

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 17:56:26 ; Search time 19 Seconds  
(Without alignments)  
75.022 Million cell updates/sec

Title: US-09-993-366-1  
Perfect score: 66  
Sequence: 1 KLVFPATELHERL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 124384

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database :

1: /published/Applications\_AA:\*  
2: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	13	9	US-09-991-809-2
2	36	54.5	38	10	US-09-864-761-42714
3	34	51.5	35	10	US-09-864-761-36887
4	34	51.5	47	10	US-09-864-761-40998
5	32	48.5	15	9	US-10-235-483-55
6	32	48.5	20	10	US-09-736-076-7
7	32	48.5	24	9	US-10-097-065-417
8	31	47.0	9	9	US-10-235-483-52
9	31	47.0	43	10	US-09-764-877-1705
10	31	47.0	47	9	US-10-062-548-135
11	30	45.5	24	9	US-10-074-475-244
12	30	45.5	35	9	US-09-809-391-337
13	30	45.5	41	10	US-09-864-761-36531
14	29	43.9	43	9	US-09-809-391-638
15	29	43.9	34	8	US-08-424-5508-406
16	29	43.9	50	9	US-09-975-143-1
17	28.5	43.2	37	9	US-10-074-475-204
18	28	42.4	23	9	US-10-007-768-7
19	28	42.4	23	10	US-09-821-984-4

20	28	42.4	29	9	US-10-023-282-1187	Sequence 1187, App
21	28	42.4	30	10	US-09-864-761-43540	Sequence 43540, A
22	28	42.4	32	10	US-10-174-410-236	Sequence 236, App
23	28	42.4	34	10	US-09-746-919-40	Sequence 40, App
24	28	42.4	40	9	US-10-002-344A-187	Sequence 187, App
25	28	42.4	42	10	US-09-864-761-47923	Sequence 47923, A
26	27	40.9	11	9	US-10-062-710-156	Sequence 156, App
27	27	40.9	12	9	US-10-254-446A-220	Sequence 220, App
28	27	40.9	14	9	US-10-166-221-7	Sequence 7, App
29	27	40.9	19	9	US-10-225-667A-1488	Sequence 1488, App
30	27	40.9	29	9	US-09-974-879-527	Sequence 527, App
31	27	40.9	29	9	US-09-305-736-529	Sequence 529, App
32	27	40.9	45	10	US-09-864-761-47137	Sequence 47137, A
33	26	39.4	5	9	US-10-235-483-32	Sequence 32, App
34	26	39.4	15	9	US-10-211-069-18	Sequence 18, App
35	26	39.4	19	9	US-10-218-743-11	Sequence 11, App
36	26	39.4	20	9	US-10-218-743-3	Sequence 3, App
37	26	39.4	28	10	US-09-864-761-39493	Sequence 39493, A
38	26	39.4	28	10	US-09-864-761-40300	Sequence 40300, A
39	26	39.4	31	10	US-09-864-761-35615	Sequence 35615, A
40	26	39.4	32	9	US-10-174-410-240	Sequence 240, App
41	26	39.4	32	10	US-09-864-761-37322	Sequence 37322, A
42	26	39.4	33	10	US-09-864-761-43699	Sequence 43699, A
43	26	39.4	35	9	US-10-029-217A-21	Sequence 21, App
44	26	39.4	38	10	US-09-864-761-43123	Sequence 43123, A
45	26	39.4	39	10	US-09-925-301-1285	Sequence 1285, App
46	26	39.4	41	10	US-09-864-761-38257	Sequence 38257, A
47	26	39.4	42	9	US-10-016-157A-186	Sequence 186, App
48	26	39.4	48	10	US-09-941-314-17	Sequence 17, App
49	25	37.9	12	9	US-09-826-290-423	Sequence 423, App
50	25	37.9	18	9	US-10-181-654-11	Sequence 11, App
51	25	37.9	22	10	US-09-864-761-42734	Sequence 42734, A
52	25	37.9	22	10	US-09-864-761-46463	Sequence 46463, A
53	25	37.9	24	9	US-10-023-282-1153	Sequence 1153, App
54	25	37.9	25	10	US-09-798-831-1	Sequence 1, App
55	25	37.9	25	10	US-09-798-831-2	Sequence 2, App
56	25	37.9	25	10	US-09-798-831-5	Sequence 5, App
57	25	37.9	25	10	US-09-798-831-6	Sequence 6, App
58	25	37.9	25	10	US-09-798-831-7	Sequence 7, App
59	25	37.9	25	12	US-10-066-895-17	Sequence 17, App
60	25	37.9	27	9	US-10-181-654-42	Sequence 32, App
61	25	37.9	32	9	US-10-174-410-247	Sequence 247, App
62	25	37.9	34	10	US-09-864-761-34126	Sequence 34126, A
63	25	37.9	34	10	US-09-864-761-47026	Sequence 47026, A
64	25	37.9	35	9	US-10-029-217A-8	Sequence 8, App
65	25	37.9	35	9	US-10-029-217A-9	Sequence 9, App
66	25	37.9	37	9	US-10-091-504-777	Sequence 777, App
67	25	37.9	37	10	US-09-764-869-777	Sequence 777, App
68	25	37.9	38	9	US-10-083-357-1067	Sequence 1067, App
69	25	37.9	39	10	US-09-864-761-36442	Sequence 36442, A
70	25	37.9	40	9	US-09-989-919-87	Sequence 87, App
71	25	37.9	41	9	US-10-000-256A-137	Sequence 137, App
72	25	37.9	43	10	US-09-864-761-42201	Sequence 42201, A
73	25	37.9	46	10	US-09-922-261-50	Sequence 50, App
74	25	37.9	48	10	US-09-864-761-47284	Sequence 47284, A
75	25	37.9	49	9	US-10-106-698-6938	Sequence 6938, App

#### ALIGNMENTS

US-09-991-809-2  
Sequence 2, Application US/09991809  
Publication No. US20030100014A1  
GENERAL INFORMATION:  
APPLICANT: Jackowski, George  
TITLE OF INVENTION: Apolipoprotein Biopolymer Markers Predictive of Type II Diabet  
FILE REFERENCE: 2132.111  
CURRENT APPLICATION NUMBER: US/09/991,809  
CURRENT FILING DATE: 2001-11-23  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.1

SEQ ID NO 2  
LENGTH: 13  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-991-809-2

Query Match 100.0%; Score 66; DB 9; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVFATLHERL 13  
Db 1 KLVFATLHERL 13

RESULT 2  
US-09-864-761-42714  
Sequence 42714, Application US/09864761

Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeonlca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 42714  
LENGTH: 38  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC009408.2

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.96  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1  
OTHER INFORMATION: EST-HUMAN HIT: AA081792.1, EVALUE 2.90e-01  
US-09-864-761-42714  
Query Match 54.5%; Score 36; DB 10; Length 38;  
Best Local Similarity 61.5%; Pred. No. 9.3;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVFATLHERL 13  
Db 22 KLVFATLHERL 34

RESULT 3  
US-09-864-761-36887  
Sequence 36887, Application US/09864761

Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
FILE REFERENCE: Aeonlca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 36887  
LENGTH: 35  
TYPE: PRT

ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC008085.1  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.82  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.81  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.92  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.68  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.87  
US-09-864-761-36887

Query Match 51.5% Score 34; DB 10; Length 35;  
Best Local Similarity 72.7% Pred. No. 20;  
Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

Qy 1 KLVPFATELHE 11  
Db 13 KMVP-TELHE 21

RESULT 4  
US-09-864-761-40998  
Sequence 40998, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aecomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864, 761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
SEQ ID NO 40998  
LENGTH: 47  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC018742.2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8  
OTHER INFORMATION: EST\_HUMAN HIT: R31169.1, EVALU2.30e-02  
OTHER INFORMATION: SWISSPROT HIT: O46409, EVALU1.00e-02  
US-09-864-761-40998

Query Match 51.5% Score 34; DB 10; Length 47;  
Best Local Similarity 70.0% Pred. No. 27;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVPFATELH 10  
Db 7 KLVPFIMELN 16

RESULT 5  
US-10-235-483-55  
Sequence 55, Application US/10235483  
Publication No. US20030087407A1  
GENERAL INFORMATION:  
APPLICANT: SOTO-JARA, Claudio  
BAUMANN, Marc  
FRANGIONE, Blas  
TITLE OF INVENTION: PEPTIDES AND PHARMACEUTICAL  
COMPOSITIONS THEREOF FOR TREATMENT OF DISORDERS OR DISEA  
ASSOCIATED WITH PROTEIN FOLDING INTO AMYLOID OR AMYLOID-  
DEPOSITS  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESS: BROWDY AND NEWMARK  
STREET: 419 Seventh Street, N.W., Suite 400  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/235,483  
FILING DATE: 06-Sep-2002  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/766,596  
FILING DATE: <unknown>  
APPLICATION NUMBER: US 08/630,645  
FILING DATE: 10-APR-1996  
APPLICATION NUMBER: US 08/478,326  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: YUN, Allen C.  
REGISTRATION NUMBER: 37,971  
REFERENCE/DOCKET NUMBER: SOTO-JARA-1A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 55:

SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRADEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 55:  
US-10-235-483-55

Query Match 48.5%; Score 32; DB 9; Length 15;  
Best Local Similarity 66.7%; Pred. NO. 18;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVPEATEL 9  
DB 5 KLVPEADVD 13

RESULT 6  
US-09-736-076-7  
Sequence 7, Application US/09736076  
Patent No. US20020049301A1  
GENERAL INFORMATION:  
APPLICANT: Ben-Sasson Shmuel A.  
TITLE OF INVENTION: SHORT PEPTIDES WHICH SELECTIVELY  
MODULATE THE ACTIVITY OF SERINE/THREONINE KINASES  
FILE REFERENCE: 1242.1015-009  
CURRENT APPLICATION NUMBER: US/09/736,076  
CURRENT FILING DATE: 2000-12-13  
PRIOR APPLICATION NUMBER: US 08/861,338  
PRIOR FILING DATE: 1997-05-21  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 20  
TYPE: PRT  
ORGANISM: Unknown  
FEATURE:  
OTHER INFORMATION: AKT/PKB  
US-09-736-076-7

Query Match 48.5%; Score 32; DB 10; Length 20;  
Best Local Similarity 54.5%; Pred. NO. 24;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 VPFAETLHERL 13  
DB 8 LPFYNDQHERL 18

RESULT 7  
US-10-097-065-417  
Sequence 417, Application US/10097065  
Publication No. US20030055236A1  
GENERAL INFORMATION:  
APPLICANT: Moore, Paul A. et al.  
TITLE OF INVENTION: 110 Human Secreted Proteins  
FILE REFERENCE: P2021P1  
CURRENT APPLICATION NUMBER: US/10/097,065  
CURRENT FILING DATE: 2002-03-14  
PRIOR APPLICATION NUMBER: PCT/US98/27059  
PRIOR FILING DATE: 1998-12-17  
PRIOR APPLICATION NUMBER: 60/070,923  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/068,007  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/068,057  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/068,006  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/068,369  
PRIOR FILING DATE: 1997-12-19  
PRIOR APPLICATION NUMBER: 60/068,367

PRIOR FILING DATE: 1997-12-19  
PRIOR APPLICATION NUMBER: 60/068,368  
PRIOR FILING DATE: 1997-12-19  
PRIOR APPLICATION NUMBER: 60/068,169  
PRIOR FILING DATE: 1997-12-19  
PRIOR APPLICATION NUMBER: 60/068,053  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/068,064  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/068,054  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/068,008  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/068,365  
PRIOR FILING DATE: 1997-12-19  
NUMBER OF SEQ ID NOS: 672  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 417  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-097-065-417

Query Match 48.5%; Score 32; DB 9; Length 24;  
Best Local Similarity 53.8%; Pred. NO. 30;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPEATELHERL 13  
DB 12 KVPASQLEFVRL 24

RESULT 8  
US-10-235-483-52  
Sequence 52, Application US/10235483  
Publication No. US20030087407A1  
GENERAL INFORMATION:  
APPLICANT: SOTO-JARA, Claudio  
BAUMANN, Marc  
FRANCIONE, Blas  
TITLE OF INVENTION: PEPTIDES AND PHARMACEUTICAL  
COMPOSITIONS THEREOF FOR TREATMENT OF DISORDERS OR DISEA  
ASSOCIATED WITH PROTEIN FOLDING INTO AMYLOID OR AMYLOID-  
DEPOSITS  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROMDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 400  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/235,483  
FILING DATE: 06-Sep-2002  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/766,596  
FILING DATE: <unknown>  
APPLICATION NUMBER: US 08/630,645  
FILING DATE: 10-APR-1996  
APPLICATION NUMBER: US 08/478,326  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: YUN, Allen C  
REGISTRATION NUMBER: 37,971  
REFERENCE/DOCKET NUMBER: SOTO-JARA-1A  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 52:  
US-10-235-483-52

Query Match  
Best Local Similarity 47.0%; Score 31; DB 9; Length 9;  
Best Local Similarity 75.0%; Pred. No. 3.8e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVFATE 8  
111111  
DB 2 KLVFAD 9

RESULT 9  
US-09-764-877-1705  
Sequence 1705, Application US/09764877  
Patent No. US20020147140A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC005  
CURRENT APPLICATION NUMBER: US/09/764,877  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 4031  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 1705  
LENGTH: 43  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-877-1705

Query Match  
Best Local Similarity 47.0%; Score 31; DB 10; Length 43;  
Best Local Similarity 50.0%; Pred. No. 86;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 VFPATELH 12  
111111  
DB 10 VFPLEHRSK 19

RESULT 10  
US-10-062-548-135  
Sequence 135, Application US/10062548  
Publication No. US20030096982A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 44 Human Secreted Proteins  
FILE REFERENCE: P2024P1  
CURRENT APPLICATION NUMBER: US/10/062,548  
CURRENT FILING DATE: 2002-02-05  
Prior Application Number: 09/369,247  
Prior Filing Date: 1998-08-05  
Prior Application Number: 60/074,118  
Prior Filing Date: 1998-02-09  
Prior Application Number: 60/074,157  
Prior Filing Date: 1998-02-09  
Prior Application Number: 60/074,137  
Prior Filing Date: 1998-02-09  
Prior Application Number: 60/074,341  
Prior Filing Date: 1998-02-09  
Prior Application Number: 60/074,141  
Prior Filing Date: 1998-02-09  
NUMBER OF SEQ ID NOS: 172  
SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 135  
LENGTH: 47  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-062-548-135

Query Match  
Best Local Similarity 47.0%; Score 31; DB 9; Length 47;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ELHERL 13  
111111  
DB 8 ELHERL 13

RESULT 11  
US-10-074-475-244  
Sequence 244, Application US/10074475  
Publication No. US20030092898A1  
GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Hu, Ping  
APPLICANT: Recipon, Herve  
APPLICANT: Katta, Kalpana  
APPLICANT: Cafterkey, Robert  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific  
FILE REFERENCE: DEX-0313  
CURRENT APPLICATION NUMBER: US/10/074,475  
CURRENT FILING DATE: 2002-02-13  
Prior Application Number: 60/268,292  
Prior Filing Date: 2001-02-13  
NUMBER OF SEQ ID NOS: 295  
SOFTWARE: Patentln version 3.1  
SEQ ID NO 244  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-074-475-244

Query Match  
Best Local Similarity 45.5%; Score 30; DB 9; Length 24;  
Best Local Similarity 62.5%; Pred. No. 69;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LVFPATEL 9  
111111  
DB 10 LVFPAGEI 17

RESULT 12  
US-09-809-391-337  
Sequence 337, Application US/09809391  
Publication No. US20030049618A1  
GENERAL INFORMATION:  
APPLICANT: Ruden et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P2  
CURRENT APPLICATION NUMBER: US/09/809,391  
CURRENT FILING DATE: 2001-03-16  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 761  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 337  
LENGTH: 35  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-809-391-337

Query Match  
Best Local Similarity 45.5%; Score 30; DB 9; Length 35;  
Best Local Similarity 41.7%; Pred. No. 1e+02;

Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 LVFPAELHERL 13  
 Db 18 ILPVCALHHEEL 29

RESULT 13  
 US-09-864-761-36531  
 ; Sequence 36531, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharon G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 ; FILE REFERENCE: Aeomica-X-1  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 36531  
 ; LENGTH: 41  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AF067844.1  
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL ~ 2.9  
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL ~ 4.5  
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL ~ 1.1  
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL ~ 1.1  
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL ~ 5.6  
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL ~ 2  
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL ~ 10  
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL ~ 3.1

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL ~ 4.9  
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL ~ 8.2  
 ; OTHER INFORMATION: EST\_HUMAN HIT: A1589294.1, EVALU 5.00e-17  
 US-09-864-761-36531

Query Match 45.5%; Score 30; DB 10; Length 41;  
 Best Local Similarity 58.3%; Prod. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LVFPAELHERL 13  
 Db 12 LAPLATEMKRL 23

RESULT 14  
 US-09-809-391-638  
 ; Sequence 638, Application US/09809391  
 ; Publication No. US20030049618A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruben et al.  
 ; TITLE OF INVENTION: 186 Human Secreted proteins  
 ; FILE REFERENCE: P2002P2  
 ; CURRENT FILING DATE: 2001-03-16  
 ; PRIOR APPLICATION NUMBER: US/09/809,391  
 ; NUMBER OF SEQ ID NOS: 761  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 638  
 ; LENGTH: 43  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-809-391-638

Query Match 45.5%; Score 30; DB 9; Length 43;  
 Best Local Similarity 41.7%; Prod. No. 1.3e+02;  
 Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 LVFPAELHERL 13  
 Db 30 ILPVCALHHEEL 41

RESULT 15  
 US-08-424-5508-406  
 ; Sequence 406, Application US/08424550B  
 ; Patent No. US20020119447A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN N. SIMONS  
 ; APPLICANT: TAM J. PILOT-MATIAS  
 ; APPLICANT: GEORGE J. DAWSON  
 ; APPLICANT: GEORGE G. SCHLAUDER  
 ; APPLICANT: SURESH M. DESAI  
 ; APPLICANT: THOMAS P. LEARY  
 ; APPLICANT: ANTHONY SCOTT MUEHROFF  
 ; APPLICANT: JAMES C. ERKER  
 ; APPLICANT: SHERI L. BUIK  
 ; APPLICANT: ISA K. MUSHAWAR  
 ; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS  
 ; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE  
 ; NUMBER OF SEQUENCES: 716  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D  
 ; STREET: 100 ABBOTT PARK ROAD  
 ; CITY: ABBOTT PARK  
 ; STATE: IL  
 ; COUNTRY: USA  
 ; ZIP: 60064-3500  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/424,550B  
FILING DATE:  
CLASSIFICATION: 435435  
ATTORNEY/AGENT INFORMATION:  
NAME: POREMBSKI, PRISCILLA E.  
REGISTRATION NUMBER: 33,207  
REFERENCE/DOCKET NUMBER: 5527.PC.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708-937-6365  
TELEFAX: 708-938-2623  
INFORMATION FOR SEQ ID NO: 406:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-424-550B-406

Query Match 43.9%; Score 29; DB 8; Length 34;  
Best Local Similarity 45.5%; Pred. No. 1.5e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 LVFPATELHER 12  
|:|:|:|:|  
Db 13 LLPMANDAHRV 23

RESULT 16  
US-09-975-143-1  
Sequence 1, Application US/09975143  
Patent No. US20020155513A1  
GENERAL INFORMATION:  
APPLICANT: HSU, Daniel, K.  
APPLICANT: LIU, Fu-Tong  
APPLICANT: DOMLING, Christopher, A.  
TITLE OF INVENTION: GALECTIN EXPRESSION IS INDUCED IN  
TITLE OF INVENTION: CIRRHOTIC LIVER AND HEPATOCELLULAR CARCINOMA  
FILE REFERENCE: DANHSU.001C1  
CURRENT APPLICATION NUMBER: US/09/975,143  
CURRENT FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: PCT/US00/08561  
PRIOR FILING DATE: 2000-03-29  
NUMBER OF SEQ ID NOS: 47  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 50  
TYPE: PRT  
ORGANISM: chicken  
US-09-975-143-1

Query Match 43.9%; Score 29; DB 9; Length 50;  
Best Local Similarity 41.7%; Pred. No. 2.3e+02;  
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 LVFPATELHERL 13  
|:|:|:|:|:|  
Db 17 LLPPLSSPLHPQL 28

RESULT 17  
US-10-074-475-204  
Sequence 204, Application US/100744475  
Publication No. US20030092898A1  
GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Hu, Ping  
APPLICANT: Reclapon, Hervé  
APPLICANT: Karra, Kalpana  
APPLICANT: Cafierkey, Robert  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific

TITLE OF INVENTION: Genes and Proteins  
FILE REFERENCE: DEX-0313  
CURRENT APPLICATION NUMBER: US/10/074,475  
CURRENT FILING DATE: 2002-02-13  
PRIOR APPLICATION NUMBER: 60/268,292  
PRIOR FILING DATE: 2001-02-13  
NUMBER OF SEQ ID NOS: 295  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 204  
LENGTH: 37  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-074-475-204

Query Match 43.2%; Score 28.5; DB 9; Length 37;  
Best Local Similarity 61.5%; Pred. No. 2.1e+02;  
Matches 8; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

OY 1 KLVPATELHERL 13  
|:|:|:|:|:|  
Db 22 KLRP-ASKLHERL 33

RESULT 18  
US-10-007-768-7  
Sequence 7, Application US/10007768  
Publication No. US20030099929A1  
GENERAL INFORMATION:  
APPLICANT: Vojdani, Aristo  
TITLE OF INVENTION: SALIVA IMMUNOASSAY FOR DETECTION OF  
TITLE OF INVENTION: EXPOSURE TO INFECTIOUS DISEASES  
FILE REFERENCE: IMSC12.003A  
CURRENT APPLICATION NUMBER: US/10/007,768  
CURRENT FILING DATE: 2002-07-02  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Coxsackie virus peptide  
US-10-007-768-7

Query Match 42.4%; Score 28; DB 9; Length 23;  
Best Local Similarity 46.2%; Pred. No. 1.5e+02;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 KLVPATELHERL 13  
|:|:|:|:|:|  
Db 8 KILPEVKEKHEFL 20

RESULT 19  
US-09-821-984-4  
Sequence 4, Application US/09821984  
Patent No. US20020004205A1  
GENERAL INFORMATION:  
APPLICANT: Consler, Thomas G.  
APPLICANT: Iannone, Marie A.  
APPLICANT: Gray, John G.  
APPLICANT: Stimmel, Julia E.  
TITLE OF INVENTION: METHOD OF INVESTIGATING FUNCTIONAL  
TITLE OF INVENTION: MOLECULAR INTERACTIONS AND REAGENTS FOR USE THEREIN  
FILE REFERENCE: 07083.000702  
CURRENT APPLICATION NUMBER: US/09/821,984  
CURRENT FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: 60/193,826  
PRIOR FILING DATE: 2000-03-31  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Artificial Sequence

FEATURE:	DESCRIPTION: Description of Artificial Sequence:/note
OTHER INFORMATION:	synthetic construct
US-09-821-984-4	
Query Match	42.4% Score 28; DB 10; Length 23;
Best Local Similarity	50.0%; Pred. No. 1.5e+02;
Matches	6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY	2 LVPFATLHERL 13
Db	2 LVPDAASKHKL 13
RESULT 20	
US-10-023-282-1187	
Sequence 1187, Application US/10023282	
Publication No. US20030092893A1	
GENERAL INFORMATION:	
APPLICANT: Young et al.	
TITLE OF INVENTION: 207 Human Secreted Proteins	
FILE REFERENCE: P2007P1	
CURRENT APPLICATION NUMBER: US/10/023,282	
EARLIER FILING DATE: 2001-12-20	
EARLIER FILING DATE: 1998-12-04	
EARLIER APPLICATION NUMBER: PCT/US98/11422	
EARLIER FILING DATE: 1998-06-04	
EARLIER APPLICATION NUMBER: 60/048,885	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/049,375	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,881	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,880	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,896	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/049,020	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,876	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,895	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,884	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,894	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,971	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,964	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,882	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,899	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,893	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,900	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,901	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,892	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,915	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/049,019	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,970	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,972	
EARLIER FILING DATE: 1997-06-06	
EARLIER APPLICATION NUMBER: 60/048,916	

```

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.20
SEQ ID NO 1187
LENGTH: 29
TYPE: PRT
ORGANISM: Homo sapiens
US-10-023-282-1187

Query Match          42.4%  Score 28;  DB 9;  Length 29;
Best Local Similarity 46.2%  Pred. NO. 1.9e+02;
Matches      6;  Conservatively      1;  Mismatches      6;  Indels      0;  Gaps      0;

QY      1  KLVFPATLHERL 13
      1 1 1 1 1 1
Db      11 KVFPLLLHEEL 23

RESULT 21
US-09-864-761-43540
; Sequence 43540, Application US/09864761
; Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359

```



;; ORGANISM: Bacillus subtilis  
US-10-174-410-236  
Query Match 42.4%; Score 28; DB 10; Length 32;  
Best Local Similarity 23.1%; Pred. No. 2.2e+02;  
Matches 3; Conservative 7; Mismatches 3; Indels 0; Gaps 0;  
OY 1 KLVPPATELHERL 13  
1 : : : : :  
Db 14 KSOPMPSLHPKM 26  
1 : : : : :  
RESULT 22  
US-10-174-410-236  
; Sequence 236, Application US/10174410  
; Publication No. US20030073134A1  
; GENERAL INFORMATION:  
; APPLICANT: Louie, Gordon V.  
; APPLICANT: Buchanan, Sean Grant  
; APPLICANT: Gajiwala, Ketan S.  
; APPLICANT: Sauder, J. Michael  
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF  
; FILE REFERENCE: 524982000300  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US/10/174,410  
; PRIOR FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 60/299,058  
; NUMBER OF SEQ ID NOS: 336  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 236  
; LENGTH: 32  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC005703.1  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.71  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.81  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.94  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.82  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.83  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
US-09-864-761-43540  
Query Match 42.4%; Score 28; DB 10; Length 30;  
Best Local Similarity 38.5%; Pred. No. 2e+02;  
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
OY 1 KLVPPATELHERL 13  
1 : : : : :  
Db 14 KSOPMPSLHPKM 26  
1 : : : : :  
RESULT 22  
US-10-174-410-236  
; Sequence 236, Application US/10174410  
; Publication No. US20030073134A1  
; GENERAL INFORMATION:  
; APPLICANT: Louie, Gordon V.  
; APPLICANT: Buchanan, Sean Grant  
; APPLICANT: Gajiwala, Ketan S.  
; APPLICANT: Sauder, J. Michael  
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF  
; FILE REFERENCE: 524982000300  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US/10/174,410  
; PRIOR FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 60/299,058  
; NUMBER OF SEQ ID NOS: 336  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 236  
; LENGTH: 32  
; TYPE: PRT

;; ORGANISM: Bacillus subtilis  
US-10-174-410-236  
Query Match 42.4%; Score 28; DB 9; Length 32;  
Best Local Similarity 23.1%; Pred. No. 2.2e+02;  
Matches 3; Conservative 7; Mismatches 3; Indels 0; Gaps 0;  
OY 1 KLVPPATELHERL 13  
1 : : : : :  
Db 2 KMLPTIEDMRKRI 14  
1 : : : : :  
RESULT 23  
US-09-746-919-40  
; Sequence 40, Application US/09746919  
; Patent No. US20020013452A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Howard M.  
; APPLICANT: Pontzer, Carol H.  
; TITLE OF INVENTION: Interferon Tau Compositions and  
; TITLE OF INVENTION: Methods of Use  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Ave., Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentia Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/746,919  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/045,467  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/438,753  
FILING DATE: 10-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/139,891  
FILING DATE: 19-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847,741  
FILING DATE: 09-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/318,050  
FILING DATE: 02-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/969,890  
FILING DATE: 30-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Dehlinger, Peter J.  
REGISTRATION NUMBER: 28,006  
REFERENCE/DOCKET NUMBER: 5600-0001.36  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-0880  
TELEFAX: 650-324-0960  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Amino acid sequence of fragment  
INDIVIDUAL ISOLATE: 139-172 of SEQ ID NO:33  
US-09-746-919-40

Query Match 42.4%; Score 28; DB 10; Length 34;  
Best Local Similarity 62.5%; Pred. No. 2.3e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 ATEUHERL 13  
: | | | | |  
DB 17 STSLHKRL 24

## RESULT 24

US-10-002-344A-187  
; Sequence 187, Application US/10002344A  
; Patent No. US20020172959A1  
; GENERAL INFORMATION:  
; APPLICANT: Recipon, Herve  
; APPLICANT: Sun, Yongming  
; APPLICANT: Chen, Sei-yu  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Turner, Leah  
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot  
; FILE REFERENCE: DEX-0241  
; CURRENT APPLICATION NUMBER: US/10/002,344A  
; CURRENT FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 60/242,998  
; PRIOR FILING DATE: 2000-10-25  
; NUMBER OF SEQ ID NOS: 277  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 187  
; LENGTH: 40  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-002-344A-187

Query Match 42.4%; Score 28; DB 9; Length 40;  
Best Local Similarity 60.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LVFPATELHE 11  
: | | | | |  
DB 15 LMSFAGESHE 24

## RESULT 25

US-09-864-761-47923  
; Sequence 47923, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecm1ca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47923  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens

; OTHER INFORMATION: MAP TO AC009493.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
; OTHER INFORMATION: EST\_HUMAN HIT: R13658.1, EVALU0 3.20e-01  
US-09-864-761-47923

Query Match 42.4%; Score 28; DB 10; Length 42;  
Best Local Similarity 50.0%; Pred. No. 2.9e+02;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VFPATELH 10  
: | | | | |  
DB 22 IFFGMEFH 29

## RESULT 26

US-10-062-710-156  
; Sequence 156, Application US/10062710  
; Publication No. US20030049253A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Frank O.  
; APPLICANT: Chu, Yong-Liang  
; APPLICANT: Qiu, Jian-Tai  
; TITLE OF INVENTION: Polymeric Conjugates for Delivery of  
; TITLE OF INVENTION: MHC-Recognized Epitopes  
; FILE REFERENCE: 3781-001-27  
; CURRENT APPLICATION NUMBER: US/10/062,710  
; CURRENT FILING DATE: 2002-02-05  
; PRIOR APPLICATION NUMBER: US 60/310,498  
; PRIOR FILING DATE: 2001-08-08  
; NUMBER OF SEQ ID NOS: 232  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 156  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HIV B Cell Epitopes  
US-10-062-710-156

Query Match 40.9%; Score 27; DB 9; Length 11;  
Best Local Similarity 57.1%; Pred. No. 1e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 PFATELH 10

Db 1 PYDTEVH 7

RESULT 27

US-10-254-446A-220

Sequence 220, Application US/10254446A

Publication No. US20030113714A1

GENERAL INFORMATION:

APPLICANT: Belcher, Angela M

APPLICANT: Smalley, Richard E.

APPLICANT: Ryan, Esther

APPLICANT: Lee, Seung-Wuk

TITLE OF INVENTION: BIOLOGICAL CONTROL OF NANOPARTICLES

FILE REFERENCE: 119927-1066

CURRENT APPLICATION NUMBER: US/10/254,446A

CURRENT FILING DATE: 2003-02-19

PRIOR APPLICATION NUMBER: 60/325,664

PRIOR FILING DATE: 2001-09-28

NUMBER OF SEQ ID NOS: 245

SOFTWARE: PatentIn version 3.1

SEQ ID NO 220

LENGTH: 12

TYPE: PRT

ORGANISM: artificial sequence

FEATURE:

OTHER INFORMATION: peptide with peptide binding sequence retrieved from phage biopan

US-10-254-446A-220

Query Match

Best Local Similarity 40.9%; Score 27; DB 9; Length 12;

Best Local Similarity 57.1%; Pred. No. 1.1e+02;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LVPFAT 8

Db 5 LIPFASQ 11

RESULT 28

US-10-166-221-7

Sequence 7, Application US/10166221

Publication No. US20030022211A1

GENERAL INFORMATION:

APPLICANT: Wyeich

TITLE OF INVENTION: Identification of a novel GPCR and uses therefor

FILE REFERENCE: AM100539

CURRENT APPLICATION NUMBER: US/10/166,221

CURRENT FILING DATE: 2002-06-07

NUMBER OF SEQ ID NOS: 7

SOFTWARE: PatentIn version 3.1

SEQ ID NO 7

LENGTH: 14

TYPE: PRT

ORGANISM: homo sapiens

US-10-166-221-7

Query Match

Best Local Similarity 40.9%; Score 27; DB 9; Length 14;

Best Local Similarity 71.4%; Pred. No. 1.3e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVPFAT 7

Db 4 ELVPFVT 10

RESULT 29

US-10-225-567A-1488

Sequence 1488, Application US/10225567A

Publication No. US20030113798A1

GENERAL INFORMATION:

APPLICANT: Lifespan Biosciences

APPLICANT: Brown, Joseph P.

APPLICANT: Burner, Glenna C.

APPLICANT: Roush, Christine L.

TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPT

FILE REFERENCE: 1920-4-4

CURRENT APPLICATION NUMBER: US/10/225,567A

CURRENT FILING DATE: 2001-12-19

PRIOR APPLICATION NUMBER: 60/257,144

PRIOR FILING DATE: 2000-12-19

NUMBER OF SEQ ID NOS: 2292

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1488

LENGTH: 19

TYPE: PRT

ORGANISM: Homo sapiens

US-10-225-567A-1488

Query Match

Best Local Similarity 40.9%; Score 27; DB 9; Length 19;

Best Local Similarity 62.5%; Pred. No. 1.9e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ATELHERL 13

Db 8 AVOLHRL 15

RESULT 30

US-09-974-879-527

Sequence 527, Application US/09974879

Publication No. US20030028003A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 125 Human Secreted Proteins

FILE REFERENCE: P2020P2

CURRENT APPLICATION NUMBER: US/09/974,879

CURRENT FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: US 60/239,893

PRIOR FILING DATE: 2000-10-13

PRIOR APPLICATION NUMBER: US 09/818,683

PRIOR FILING DATE: 2001-03-28

PRIOR APPLICATION NUMBER: US 09/305,736

PRIOR FILING DATE: 1999-05-05

PRIOR APPLICATION NUMBER: PCT/US98/23435

PRIOR FILING DATE: 1998-11-04

PRIOR APPLICATION NUMBER: US 60/064,911

PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: US 60/064,912

PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: US 60/064,983

PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: US 60/064,900

PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: US 60/064,988

PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: US 60/064,987

PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: US 60/064,908

PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: US 60/064,984

PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: US 60/064,985

PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: US 60/066,094

PRIOR FILING DATE: 1997-11-17

PRIOR APPLICATION NUMBER: US 60/066,100

PRIOR FILING DATE: 1997-11-17

PRIOR APPLICATION NUMBER: US 60/066,089

PRIOR FILING DATE: 1997-11-17

PRIOR APPLICATION NUMBER: US 60/066,095

PRIOR FILING DATE: 1997-11-17

PRIOR APPLICATION NUMBER: US 60/066,090

PRIOR FILING DATE: 1997-11-17

NUMBER OF SEQ ID NOS: 611

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 527

; LENGTH: 29  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-974-879-527

Query Match 40.98; Score 27; DB 9; Length 29;  
Best Local Similarity 57.1%; Pred. No. 3e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 PRATELH 10  
| : | : | :  
Db 8 PYHTQLH 14

Search completed: July 1, 2003, 17:58:20  
Job time : 21 secs

Human bone marrow
Peptide #3646 encodes
Peptide #3744 encodes
Peptide #3591 encodes
Human peptide encodes
Peptide #9642 encodes
Protein #7699 encodes
Human brain express
Human bone marrow
Peptide #9979 encodes
Human peptide encodes
Peptide derived from
HJ loop peptide AM
Peptide sequence C
Fragment of human
Desmocollin-2 CAR
Interferon-alpha2
Human musculoskeletal
Human secreted protein
Novel human diagnosis
Novel human diagnosis
Maize ZmKCS1 peptid
Human secreted protein
Peptide #3297 encodes
Peptide #3322 encodes
Protein #3232 encodes
Human brain expresses
Human bone marrow
Peptide #3266 encodes
Peptide #3354 encodes
Peptide #3226 encodes
Human peptide encodes
Fragment of human
Desmocollin-2 CAR
Hepatitis GB virus
Novel human diagnosis
Gene 2 human secreted
Human secreted protein
Arabidopsis thaliana
Amino acid sequenc
Human OREX protein
Desmocollin-2 CAR
Human peptide #1847
Interferon-receptor
Interferon receptor
Biotinylated synthet
Molecular interacto
Plasmodium M. m.
Plasmodium PMP100 M
Insulin/insulin-like
Human alpha-1 antit
Mutated alpha-1 ant
Human secreted prote
Apolipoprotein A-I r
Peptide #6122 encod
Human brain expres
Human bone marrow
Peptide #6111 encod
Human peptide encod
Human protective s
Human novel foetal
Human immune/haema
Peptide #2536 encod
Human peptide encod
Propionibacterium

```

XX AC AAY42559;
XX DT 20-DEC-1999 (first entry)
XX XX
DE Apolipoprotein A-IV derived lipid oxidation suppressant peptide #7.
XX
KM Cholesterol; cardiovascular disease; heart disease; atherosclerosis;
KM lipoprotein; angina; myocardial infarction; stroke; thrombosis;
KM antioxidant; hypolipidemic; apolipoprotein.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key
FT Modified-site 15 Location/Qualifiers
FT /note="optionally has a C-terminal amide"
XX
PN W09950286-A2.
XX
PD 07-OCT-1999.
XX
PF 25-MAR-1999; 99WO-US06580.
XX
PR 31-MAR-1998; 98US-0080131.
XX
PA (UYCI-) UNIV CININNATI.
XX
PI Hul DY, Tso P;
XX
DR WPI; 1999-580739/49.
XX
PT Treating conditions associated with lipid oxidation or preventing
PT oxidation in lipid-containing food, lipid containing pharmaceuticals or
PT cosmetic or dermatological compositions.
XX
PS Claim 4; Page 53; 73pp: English.
XX
CC This sequence represents a peptide (#7) derived from apolipoprotein
CC (apo) A-IV with lipid oxidation inhibitory activity. Lipid oxidation
CC plays a role in the development of atherosclerosis, a main cause of
CC coronary heart disease. Atherosclerosis is thought to begin with local
CC injury to the arterial endothelium, followed by proliferation of
CC arterial smooth muscle cells, along with deposition of lipid and
CC accumulation of foam cells in the lesion. As the atherosclerotic plaque
CC develops, it progressively occludes more and more blood vessel and can
CC eventually lead to ischaemia or infarction. Both the plasma concentration
CC and qualitative characteristics of low density lipoproteins (LDL) are
CC risk factors in atherogenesis. Oxidation causes important changes in the
CC primary structure of the main LDL apolipoprotein, apolipoprotein B100
CC (apo B-100). These changes, by helping LDL absorption by macrophages,
CC causes the intracellular accumulation of esters of cholesterol and the
CC formation of foam cells, with subsequent development of the
CC atherosclerotic plaque. The apo A-IV derived peptides can be used for
CC inhibiting lipid oxidation. The peptides can be used for treating
CC conditions associated with lipid oxidation. For example, they can be used
CC for treating or inhibiting the progression of atherosclerosis. They can
CC also be used for preventing oxidation in lipid-containing foods,
CC lipid-containing pharmaceuticals or cosmetic or dermatological
CC compositions. As these peptides comprise specific portions of the native
CC apo A-IV protein, there should be no immunogenicity problems with their
CC administration to humans.
XX
SQ Sequence 15 AA:

```

```

Query Match 57.6%; Score 38; DB 20; Length 15;
Best Local Similarity 63.6%; Pred. No. 2.1;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 LVPFATELHER 12
   :||||| 11
DB 1 MVFPAEELKEK 11

```

```

RESULT 2
ABB28853
ID ABB28853 standard; Peptide; 38 AA.
XX
AC ABB28853:
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #1504 encoded by breast cell single exon nucleic acid probe.
XX
KM Human; microarray; single exon probe; gene expression; breast
KM disease; cancer.
XX
OS Homo sapiens.
XX
PN W0200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes.
XX
PS Claim 27; SEQ ID NO 11821; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and Br 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 38 AA:

```

```

Query Match 54.5%; Score 36; DB 22; Length 38;
Best Local Similarity 61.5%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 KLVFPAEELHERL 13
   :||||| 11
DB 22 KLVFPEVMDTHL 34

```

```
RESULT 3
AAM54808
ID AAM54808 standard; Protein: 38 AA.
XX
AC AAM54808;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 26913.
XX
KW Human: brain expressed exon; gene expression analysis; probe;
KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN W0200157275-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 26913; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 38 AA:
XX
Query Match 54.5%; Score 36; DB 22; Length 38;
Best Local Similarity 61.5%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 KLVFPFATLHERL 13
Db ||||| || |
22 KLVFPFVMEIDTHL 34
XX
RESULT 4
AAM67191
ID AAM67191 standard; Protein: 38 AA.
XX
AC AAM67191;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27497.
XX
KW Human: bone marrow expressed exon; gene expression analysis; probe;
KM microarray; cancer; leukaemia; lymphoma; myeloma.
XX
```

```
OS Homo sapiens.
XX
PN W0200157276-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 27497; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 38 AA:
XX
Query Match 54.5%; Score 36; DB 22; Length 38;
Best Local Similarity 61.5%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 KLVFPFATLHERL 13
Db ||||| || |
22 KLVFPFVMEIDTHL 34
XX
RESULT 5
AAM02779
ID AAM02779 standard; Protein: 38 AA.
XX
AC AAM02779;
XX
DT 09-OCT-2001 (first entry)
XX
DE Peptide #1461 encoded by probe for measuring breast gene expression.
XX
KW Probe: human; breast disease; breast cancer; development disorder;
KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
XX Homo sapiens.
XX
PN W0200157270-A2.
XX
PD 09-AUG-2001.
XX
PE 29-JAN-2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
```

(MOLE-) MOLECULAR DYNAMICS INC.  
 Penn SG, Hanzel DK, Chen W, Rank DR;  
 WPI; 2001-476286/51.  
 Novel single exon nucleic acid probe used to measuring gene expression  
 in a human breast -  
 Claim 27: SEQ ID No 11519; 322pp; English.  
 The present invention relates to novel single exon nucleic acid probes  
 (see AAI00010-AA110067). The present sequence is a peptide encoded by one  
 such probe. The probes are useful for measuring human gene expression in  
 a human breast sample, where the probe hybridises at high stringency to a  
 nucleic acid expressed in the human breast. The probes are useful for  
 predicting, diagnosing, grading, staging, monitoring and prognosing  
 diseases of the human breast, particularly those diseases with polygenic  
 aetiology. The diseases include: breast cancer, disorders of development,  
 inflammatory diseases of the breast, fibrocystic changes, proliferative  
 breast disease and non-carcinoma tumours.  
 Note: The sequence data for this patent did not form part of the printed  
 specification, but was obtained in electronic format directly from WPI  
 at ftp.wipo.int/pub/published\_pct\_sequences.  
 Sequence 38 AA:  
 Query Match 54.5%; Score 36; DB 22; Length 38;  
 Best Local Similarity 61.5%; Pred. No. 14;  
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0.  
 QY 1 KLVPEATELHERL 13  
 ||||| ||  
 DB 22 KLVPEVMDLTHL 34  
 RESULT 6  
 AAP60986  
 ID AAP60986 standard; Peptide: 16 AA.  
 XX AAP60986;  
 XX 08-AUG-1991 (first entry)  
 XX  
 DE Sequence of immunogenic fragment of apolipoprotein (ApoA1.  
 XX Apolipoprotein; immunoassay; antibody.  
 XX  
 OS Homo sapiens.  
 OS  
 PN M08604144-A.  
 PN  
 PD 17-JUL-1986.  
 PD  
 XX 26-DEC-1985; 85WO-US02569.  
 XX  
 PF 31-DEC-1984; 84US-0688040.  
 PR 26-DEC-1985; 85US-090584.  
 PR  
 XX (ITGE-) INT GENETIC ENG INC.  
 XX  
 PA Fareed G, Sen A;  
 PI  
 XX  
 DR WPI; 1986-196930/30.  
 DR  
 XX Peptide fragments of human apo:lipoprotein - used for producing  
 PT type-specific antibodies for immunoassay  
 PT  
 XX Claim 19; page 38; 53pp; English.  
 PS  
 XX The peptides of the invention are conjugated with carrier proteins  
 CC and used to produce type-specific, non-cross-reactive antibodies by  
 CC

CC	immunisation. The antibodies may then be used in immunoassays to
CC	identify and quantitate specific APLs. Peptide fragments without
CC	the C-terminal Cys residue are also claimed.
XX	
SQ	Sequence 16 AA:
OY	Query Match 51.5%; Score 34; DB 7; Length 16;
	Best Local Similarity 50.0%; Ped. No. 13;
Matches	6; Conservative 3; Mismatches 3; Indels 0; Gaps 0.
DB	2 LVPFATLHERL 13     : :   :   : 4 LAPSYDELRRRL 15
RESULT 7	
ABB31023	ID ABB31023 standard; Peptide: 35 AA.
XX	AC ABB31023;
XX	DT 01-FEB-2002 (first entry)
DE	Peptide #3674 encoded by breast cell single exon nucleic acid probe.
XX	
KM	Human; microarray; single exon probe; gene expression; breast;
KM	disease: cancer.
XX	
OS	Homo sapiens.
XX	
PN	WO200157271-A2.
PD	
XX	09-AUG-2001.
PE	30-JAN-2001; 2001WO-US000662.
XX	
PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
DR	
XX	WPI; 2001-496933/54.
PT	New spatially-addressable set of single exon nucleic acid probes,
PT	useful for measuring gene expression in sample derived from human
PT	breast, comprises number of single exon nucleic acid probes -
PS	
XX	Claim 27; SEQ ID NO 13991; 327bp + sequence listing; English.
CC	The invention relates to a spatially-addressable set of single exon
CC	nucleic acid probes for measuring gene expression in a sample derived
CC	from human breast and BR 474 cells. The method involves contacting
CC	the probes with a collection of detectably labeled nucleic acids
CC	derived from mRNA of human breast, and then measuring the label
CC	bound to each probe of the microarray. The probes are useful for
CC	verifying the expression of regions of genomic DNA predicted to
CC	encode proteins. They are useful for gene discovery, and for
CC	determining predisposition and/or prognosing breast disease. Gene
CC	expression analysis is useful for assessing the toxicity of chemical
CC	agents on cells. The microarray of this invention presents a far greater
CC	diversity of probes for measuring gene expression, with far less bias
CC	than expressed sequence tag microarrays. The method is suitable for
CC	rapid production of functional information from genomic sequence. The
CC	present sequence is a peptide encoded by a single exon nucleic acid
CC	Note: The sequence data for this patent did not form part of the



CC printed specification, but was obtained in electronic format directly  
CC from WIPo at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 35 AA;

Query Match 51.5%; Score 34; DB 22; Length 35;  
Best Local Similarity 72.7%; Pred. NO. 30;  
Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

OY 1 KLVPFATELHE 11  
I:|I |I|I|I|  
Db 13 KMP--TELHE 21

RESULT 8

ABB36212  
ID ABB36212 standard; Peptide: 35 AA.

XX ABB36212;

DT 04-FEB-2002 (first entry)

XX Peptide #3718 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

XX WO20015727-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI: 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human fetal liver.

XX Claim 27; SEQ ID NO 28847; 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC foetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPo at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 35 AA;

Query Match 51.5%; Score 34; DB 22; Length 35;  
Best Local Similarity 72.7%; Pred. NO. 30;

Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

OY 1 KLVPFATELHE 11  
I:|I |I|I|I|  
Db 13 KMP--TELHE 21

RESULT 9

ABB21589  
ID ABB21589 standard; Protein: 35 AA.

XX ABB21589;

DT 23-JAN-2002 (first entry)

XX Protein #3588 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;

XX cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00666.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI: 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX hearts -

XX Claim 15; SEQ ID NO 23359; 530pp; English.

CC The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPo  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 35 AA;

Query Match 51.5%; Score 34; DB 22; Length 35;  
Best Local Similarity 72.7%; Pred. NO. 30;

Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

OY 1 KLVPFATELHE 11  
I:|I |I|I|I|  
Db 13 KMP--TELHE 21

RESULT 10

AAM56995  
ID AAM56995 standard; Protein: 35 AA.

XX AAM56995;

XX 05-NOV-2001 (first entry)

```

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 29100.
DE
XX
XX Human: brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer.
OS
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 29100; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system,
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX
XX Sequence 35 AA;
XX
XX Query Match 51.5%; Score 34; DB 22; Length 35;
XX Best Local Similarity 72.7%; Pred. No. 30;
XX Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;
XX
XX QY 1 KLVPFATLHE 11
XX 1:11 11111
XX Db 13 KWP--TELHE 21
XX
XX RESULT 11
XX AAM69380
XX ID AAM69380 standard; Protein; 35 AA.
XX
XX AC AAM69380;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 29686.
XX
XX Human: bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX

```

```

XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 29686; 658bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention.
XX
XX Sequence 35 AA;
XX
XX Query Match 51.5%; Score 34; DB 22; Length 35;
XX Best Local Similarity 72.7%; Pred. No. 30;
XX Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;
XX
XX QY 1 KLVPFATLHE 11
XX 1:11 11111
XX Db 13 KWP--TELHE 21
XX
XX RESULT 12
XX AAM17212
XX ID AAM17212 standard; Protein; 35 AA.
XX
XX AC AAM17212;
XX
XX 12-OCT-2001 (first entry)
XX
XX Peptide #3646 encoded by probe for measuring cervical gene expression.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
XX Homo sapiens.
XX
XX WO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX

```

PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX  
PS Claim 27; SEQ ID No 22038; 487pp; English.  
XX  
CC The present invention relates to human single exon nucleic acid probes  
CC (SENP; see A110066-A128459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 35 AA:  
  
Query Match 51.5%; Score 34; DB 22; Length 35;  
Best Local Similarity 72.7%; Pred. No. 30;  
Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;  
  
QY 1 KLVPFATELHE 11  
Db 13 KMVP--TELHE 21  
  
RESULT 13  
AAM29707  
ID AAM29707 standard; Protein: 35 AA.  
XX  
AC AAM29707;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Peptide #3744 encoded by probe for measuring placental gene expression.  
XX  
KM Probe: microarray; human; placenta; antenatal diagnosis;  
KM genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD .09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00663.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488897/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta.  
XX  
PS Claim 27; SEQ ID No 29976; 654pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SENP;  
CC see A113315-A157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of

CC human genetic disorders.  
XX  
SQ Sequence 35 AA:  
  
Query Match 51.5%; Score 34; DB 22; Length 35;  
Best Local Similarity 72.7%; Pred. No. 30;  
Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;  
  
QY 1 KLVPFATELHE 11  
Db 13 KMVP--TELHE 21  
  
RESULT 14  
AAM04909  
ID AAM04909 standard; Protein: 35 AA.  
XX  
AC AAM04909;  
XX  
DT 09-OCT-2001 (first entry)  
XX  
DE Peptide #3591 encoded by probe for measuring breast gene expression.  
XX  
KM Probe: human; breast disease; breast cancer; development disorder;  
KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200157270-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 29-JAN-2001; 2001WO-US00661.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-476286/51.  
XX  
PT Novel single exon nucleic acid probe used to measuring gene expression  
PT in a human breast -  
XX  
PS Claim 27; SEQ ID No 13649; 322pp; English.  
XX  
CC The present invention relates to novel single exon nucleic acid probes  
CC (see A1100010-A110067). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for measuring human gene expression in  
CC a human breast sample, where the probe hybridises at high stringency to a  
CC nucleic acid expressed in the human breast. The probes are useful for  
CC predicting, diagnosing, grading, staging, monitoring and prognosing  
CC diseases of the human breast, particularly those diseases with polygenic  
CC aetiology. The diseases include: breast cancer, disorders of development,  
CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
CC breast disease and non-carcinoma tumours.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 35 AA:  
  
Query Match 51.5%; Score 34; DB 22; Length 35;  
Best Local Similarity 72.7%; Pred. No. 30;  
Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

Qy	I	KLVPFATELHE 11	
		I : I I I I I I	
Db	13	KMP--TELHE 21	
RESULT 15			
ABG38996			
ID	ABG38996	standard; Peptide: 35 AA.	
XX			
AC	ABG38996;		
XX			
DT	19-AUG-2002	(first entry)	
XX			
DE			
XX	Human peptide encoded by genome-derived single exon probe SEQ ID 28661.		
KW	Human; single exon probe; asthma; lung cancer; COPD; IID;		
KW	Chronic obstructive pulmonary disease; interstitial lung disease;		
KW	familial idiopathic pulmonary fibrosis; neurofibromatosis;		
KW	tuberos sclerosis; Gaucher's disease; Niemann-Pick disease;		
KW	Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;		
KW	pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;		
KW	pulmonary alveolar proteinosis; fibrocytic pulmonary dysplasia;		
KW	primary ciliary dyskinesia; pulmonary hypertension;		
KW	hyaline membrane disease.		
XX			
OS	Homo sapiens.		
XX			
PN	MO200186003-A2.		
XX			
PD	15-NOV-2001.		
XX			
PF	30-JAN-2001; 2001MO-US00665.		
XX			
PR	04-FEB-2000; 2000US-180312P.		
PR	26-MAY-2000; 2000US-207456P.		
PR	30-JUN-2000; 2000US-0608408.		
PR	03-AUG-2000; 2000US-0632366.		
PR	21-SEP-2000; 2000US-234687P.		
PR	27-SEP-2000; 2000US-236359P.		
PR	04-OCT-2000; 2000GB-0024263.		
XX			
PA	(MOE-) MOLECULAR DYNAMICS INC.		
XX			
PI	Penn SG, Hanzel DK, Chen W, Rank DR;		
DR	WPI: 2002-114183/15.		
XX			
PT	Spatially-addressable set of single exon nucleic acid probes, used to		
PT	measure gene expression in human lung samples -		
XX			
PS	Claim 27; SEQ ID No 28661; 634pp; English.		
XX			
CC	The invention relates to a spatially-addressable set of single exon		
CC	nucleic acid probes for measuring gene expression in a sample derived		
CC	from human lung comprising single exon nucleic acid probes having one of		
CC	12614 nucleic acid sequences mentioned in the specification, or their		
CC	complements or the 12387 open reading frames derived from the 12614		
CC	probes. Also included are a microarray comprising the novel set of		
CC	probes: the novel set of probes which hybridise at high stringency to a		
CC	nucleic acid expressed in the human lung; measuring gene expression in a		
CC	sample derived from human lung, comprising (a) contacting the array with		
CC	a collection of detectably labeled nucleic acids derived from human		
CC	mRNA, and (b) measuring the label detectably bound to each probe of		
CC	the array; identifying exons in a eukaryotic genome, comprising		
CC	(a) algorithmically predicting at least one exon from genomic sequences		
CC	of the eukaryote; and (b) detecting specific hybridisation of detectably		
CC	labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,		
CC	having a fragment identical to the predicted exon, the probe is included		
CC	in the above mentioned microarray; assigning exons to a single gene,		
CC	comprising (a) identifying exons from genomic sequence by the method		
CC	above and (b) measuring the expression of each of the exons in several		
CC	tissues and/or cell types using hybridisation to a single exon		
CC	microarrays having a probe with the exon, where a common pattern of		

CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO.int/pub/pub/published\_pct\_sequences.  
CC  
XX  
SQ Sequence 47 AA;

Query Match 51.5%; Score 34; DB 22; Length 47;  
Best Local Similarity 70.0%; Pred. No. 42;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KLVPFATELH 10  
||||| 11:  
Db 7 KLVPFLEMLN 16

## RESULT 17

ABB25700 standard; Protein; 47 AA.

XX ABB25700;

XX 23-JAN-2002 (first entry)

DE Protein #7699 encoded by probe for measuring heart cell gene expression.

XX Human: gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX

OS Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX hearts -

XX Claim 15; SEQ ID No 27470; 530pp; English.

CC The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC by measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 47 AA;

Query Match 51.5%; Score 34; DB 22; Length 47;  
Best Local Similarity 70.0%; Pred. No. 42;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KLVPFATELH 10  
||||| 11:  
Db 7 KLVPFLEMLN 16

## RESULT 18

AAM63020 standard; Protein; 47 AA.

XX AAM63020;

XX 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35125.

XX Human: brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX

OS Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX brains -

XX Example 4; SEQ ID NO: 35125; 650pp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.  
XX

SQ Sequence 47 AA;

Query Match 51.5%; Score 34; DB 22; Length 47;

Best Local Similarity 70.0%; Pred. No. 42;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KLVPFATELH 10  
||||| 11:  
Db 7 KLVPFLEMLN 16

## RESULT 19

AAM75831

ID AAM75831 standard; Protein; 47 AA.

XX	AA075831:
AC	
DT	06-NOV-2001 (first entry)
XX	
DE	Human bone marrow expressed probe encoded protein SEQ ID NO: 36137.
XX	
KW	Human; bone marrow expressed exon; gene expression analysis; probe;
KW	microarray; cancer; leukemia; lymphoma; myeloma.
OS	Homo sapiens.
PN	WO200157276-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US0068.
XX	
PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0633366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
DR	WPI: 2001-488900/53.
XX	
PT	Human genome-derived single exon nucleic acid probes useful for
PT	analyzing gene expression in human bone marrow -
XX	
PS	Example 4: SEQ ID NO: 36137; 658bp + Sequence Listing; English.
XX	
CC	The present invention provides a number of single exon nucleic acid
CC	probes which are derived from genomic sequences expressed in the human
CC	bone marrow. They can be used to measure gene expression in bone marrow
CC	samples, which may enable the improved diagnosis and treatment of cancers
CC	such as lymphoma, leukemia and myeloma. The present sequence is a
CC	protein encoded by one of the probes of the invention.
XX	
SQ	Sequence 47 AA:
	Query Match 51.5%; Score 34; DB 22; Length 47;
	Best Local Similarity 70.0%; Pred. No. 42;
	Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0.
QY	1 KLVPPEATELH 10
	:
Db	7 KLVFPLMELN 16
RESULT 20	
AAM35942	
ID AAM35942 standard; Protein; 47 AA.	
XX AAM35942;	
AC	
DT 17-OCT-2001 (first entry)	
XX	
DE Peptide #9979 encoded by probe for measuring placental gene expression.	
KW Probe; microarray; human; placenta; antenatal diagnosis;	
KW genetic disorder.	
XX	
OS Homo sapiens.	
XX	
PN WO200157272-A2.	
XX	
PD 09-AUG-2001.	

XX	30-JAN-2001; 2001WO-US000663.
PX	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
XX	(MOLE-) MOLECULAR DYNAMICS INC.
PA	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
DR	WPI, 2001-488897/53.
PT	Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta -
PS	Claim 27; SEQ ID No 36211; 654bp; English.
SQ	Sequence 47 AA:
Query Match	51.5%; Score 34; DB 22; Length 47; .
Best Local Similarity	70.0%; Pred. No. 42;
Matches	7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY	1 KLVPATELH 10       
Db	7 KLVPFELMELN 16
RESULT 21	
ID	ABG45288 standard; Peptide; 47 AA.
XX	ABG45288;
AC	
XX	19-AUG-2002 (first entry)
DT	
DE	Human peptide encoded by genome-derived single exon probe SEQ ID 34953.
XX	
KW	Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW	chronic obstructive pulmonary disease; interstitial lung disease;
KW	familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW	tubercous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW	Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW	pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
KW	pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW	primary ciliary dyskinesia; pulmonary hypertension;
KW	hyaline membrane disease.
XX	
OS	Homo sapiens.
XX	
PN	WO200186003-A2.
PD	
XX	15-NOV-2001.
PF	
PE	30-JAN-2001; 2001WO-US000665.
XX	
PR	04-FEB-2000; 2000US-180312P.
PR	26-MAY-2000; 2000US-207456P.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236359P.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI: 2002-114183/15.  
DR  
XX  
PT Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples -  
XX  
PS  
PS Claim 27; SEQ ID No 34953; 634pp; English.  
XX  
XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of  
CC probes; the novel set of probes which hybridise at high stringency to a  
CC nucleic acid expressed in the human lung; measuring gene expression in a  
CC sample derived from human lung, comprising (a) contacting the array with  
CC a collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of  
CC the array; identifying exons in a eukaryotic genome, comprising  
CC (a) algorithmically predicting at least one exon from genomic sequences  
CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning sequence to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene  
CC expression analysis, and for identifying exons in a gene, particularly  
CC using human lung derived mRNA and for the study of lung diseases  
CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
CC pulmonary alveolar proteinosis, Kartagener syndrome, fibrocystic  
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
CC and hyaline membrane disease. The present sequence is a peptide/protein  
CC encoded by a single exon probe of the invention.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 47 AA:  
Query Match 51.5%; Score 34; DB 23; Length 47;  
Best Local Similarity 70.0%; Pred. No. 42;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

XX  
DE Peptide derived from a cancer associated protein (FOR).  
XX  
XX  
XX Cancer associated protein; FOR gene; FRA16D; fragile site; aphidicolin;  
XX chromosomal rearrangement; cancer; splice variant; DNA instability;  
XX FRA16D oxidoreductase; neoplasia.  
XX  
XX Homo sapiens.  
XX  
XX WO200144466-A1.  
XX  
XX 21-JUN-2001.  
XX  
XX 15-DEC-2000; 2000WO-AU01539.  
XX  
XX 16-DEC-1999; 99AU-0004711.  
XX 19-APR-2000; 2000AU-0007025.  
XX  
XX (WOMEN-) WOMEN'S & CHILDREN'S HOSPITAL.  
XX  
XX Richards R, Ried K, Flumis M, Hobson L, Mangelsdorf M, Dayan S;  
XX Nancarrow J, Woolliatt E, Baker E;  
XX WPI: 2001-398151/42.  
XX  
XX Novel isolated 16q23.2 nucleic acid molecule, FRA16D oxidoreductase  
XX PT gene associated with FRA16D site, useful for early diagnosis and  
XX assessment of risk of cancers associated with the FRA16D region -  
XX  
XX Claim 36; Page 42; 150pp; English.  
XX  
XX The present sequence is derived from a cancer associated protein,  
XX encoded by the FRA16D oxidoreductase (FOR) gene. The FRA16D site is  
XX a fragile site induced by aphidicolin, which is located within the FOR  
XX gene. The fragile site is the location of breakpoints of a variety of  
XX chromosomal rearrangements, and other mutations associated with cancers.  
XX The FOR protein is expressed as a number of splice variants. FOR gene  
XX polynucleotide fragments are capable of acting as specific primers or  
XX probes for detecting cancer associated variations of DNA sequence such  
XX as a point mutation or small DNA rearrangement associated with the  
XX tumour, a breakpoint of one or more chromosomal rearrangements associated  
XX with the tumour and a pause site within the FRA16 gene. FOR nucleic acid  
XX molecules are useful as markers to identify relationship between the  
XX fragile site (FRA16D) and the DNA instability in neoplasia which  
XX allows better diagnosis of cancers associated with the region.  
XX  
SQ Sequence 32 AA:  
Query Match 50.0%; Score 33; DB 22; Length 32;  
Best Local Similarity 66.7%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 FATELHERL 13  
I: ||| I  
Db 11 FSNELHRL 19

RESULT 23  
AAW74171  
ID AAW74171 standard; peptide; 20 AA.  
XX  
XX AAW74171:  
AC  
XX 05-MAY-1999 (first entry)  
DT  
XX  
XX HJ loop peptide Akt/PKB.  
DE  
XX  
XX HJ loop; serine/threonine kinase; cancer; diabetes; obesity; therapy;  
XX haemorrhagic shock; cardiovascular disease; immunosuppressive disorder;  
XX inflammatory disorder; central nervous system disease; septic shock;  
XX Parkinson's disease; hypertension.  
XX  
XX Synthetic.

```

XX PN W09853050-A2.
XX PD 26-NOV-1998.
XX PF 20-MAY-1998; 98WO-US10319.
XX PR 21-MAY-1997; 97US-0861338.
XX PA (CHIL-) CHILDRENS MEDICAL CENT.
XX PA (YISS ) YISSUM RES & DEV CO.
XX PI Ben-Sasson SA.
XX DR WPI; 1999-070142/06.
XX PS
XX PT New peptides for modulating serine/threonine kinase activity -
XX PT comprise a sequence corresponding to the HJ loop of a
XX PT serine/threonine kinase, used for treating, e.g. cancers,
XX PT inflammatory disorders or autoimmune disorders
XX PS Claim 33; Fig 3b; 70pp; English.
XX CC This sequence represents a peptide of the invention, and is a derivative
XX CC of the HJ loop of a serine/threonine kinase (STK). The peptides can be
XX CC used for the treatment of disorders caused by overactivity or
XX CC underactivity of a STK, e.g. cancer, diabetes, obesity, haemorrhagic
XX CC shock, cardiovascular diseases (e.g. thrombosis, atherosclerosis,
XX CC arteriosclerosis, cardiac hyper trophy, ischaemia, reperfusion injury and
XX CC hypertension), immunosuppressive and inflammatory disorders (e.g. asthma,
XX CC psoriasis, systemic lupus erythematosus, diabetes mellitus, suppression
XX CC of organ transplant rejection, multiple sclerosis, inflammatory bowel
XX CC disease and AIDS), central nervous system diseases (e.g. Alzheimer's
XX CC disease, stroke and trauma), septic shock, Parkinson's disease or
XX CC hypertension. The peptides can also be used to produce antibodies which
XX CC can be used to identify cells expressing the STK and to study the
XX CC intracellular distribution of the STK. In addition, the peptides can be
XX CC used to identify and quantitate ligands which bind the HJ loop of the STK
XX CC from which the peptide was derived.
XX SQ Sequence 20 AA:
XX
XX Query Match 48.5%; Score 32; DB 20; Length 20;
XX Best Local Similarity 54.5%; Pred. No. 39;
XX Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 3 VPFA TELHERL 13
DB 8 L P F Y N Q D H E R L 18

```

```

XX PF 13-DEC-2000; 2000US-0736076.
XX PR 21-MAY-1997; 97US-0861338.
XX PA (CHIL-) CHILDRENS MEDICAL CENT.
XX PI Ben-Sasson SA.
XX DR WPI; 2002-462787/49.
XX PT New peptide from the HJ loop of serine-threonine kinase, useful for
XX PT treating e.g. cancer and for producing diagnostic antibodies -
XX PS Claim 26; Fig 3; 41pp; English.
XX CC The present invention relates to new peptides derived from the HJ loop
XX CC of a serine/threonine kinase (STK). The peptides of the invention are
XX CC used to modulate STK activity, especially for treating cancer, diabetes,
XX CC obesity or a wide variety of central nervous system, inflammatory,
XX CC autoimmune or cardiovascular disorders, also haemorrhagic shock, and for
XX CC regulating lipid metabolism. The peptides are also used to generate
XX CC antibodies that bind specifically to the parent STK, used e.g. for
XX CC identifying STK-expressing cells and to study intracellular distribution
XX CC of STK, and to identify or quantify ligands that bind to the HJ loop.
XX CC The present amino acid sequence represents the peptide sequence of the HJ
XX CC loop of serine-threonine kinase Akt/PKB. This sequence is one of the
XX CC short peptides of the invention that selectively modulate the activity
XX CC of STK.
XX SQ Sequence 20 AA:
XX
XX Query Match 48.5%; Score 32; DB 23; Length 20;
XX Best Local Similarity 54.5%; Pred. No. 39;
XX Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 3 VPFA TELHERL 13
DB 8 L P F Y N Q D H E R L 18

```

```

RESULT 24
AAU98307
ID AAU98307 standard; Peptide: 20 AA.
XX AC AAU98307;
XX DT 13-AUG-2002 (first entry)
XX DE Peptide sequence of HJ loop of serine-threonine kinase Akt/PKB.
XX KW HJ loop; serine/threonine kinase; STK; cancer; diabetes; obesity;
XX KW central nervous system disorder; inflammatory disorder;
XX KW autoimmune disorder; cardiovascular disorder; haemorrhagic shock;
XX KW lipid metabolism; cytosstatic; antidiabetic; anorectic; neurological;
XX KW antiinflammatory; immunosuppressive; cardiant; haemostatic;
XX KW modulating STK activity; Akt/PKB.
XX OS Unidentified.
XX PN US2002049301-A1.
XX PD 25-APR-2002.

```

```

RESULT 25
AAV36397
ID AAV36397 standard; Protein: 24 AA.
XX AC AAV36397;
XX DT 17-SEP-1999 (first entry)
XX DE Fragment of human secreted protein encoded by gene 5.
XX KW Human; secreted protein; cancer; tumour; developmental abnormality;
XX KW foetal deficiency; blood disorder; immune system disorder; inflammation;
XX KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
XX KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
XX KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
XX KW digestive disorder; endocrine disorder; infection; AIDS.
XX OS Homo sapiens.
XX PN W09931117-A1.
XX PD 24-JUN-1999.
XX PF 17-DEC-1998; 98WO-US27059.
XX PR 19-DEC-1997; 97US-0068369.
XX PR 18-DEC-1997; 97US-0068006.
XX PR 18-DEC-1997; 97US-0068007.
XX PR 18-DEC-1997; 97US-0068008.
XX PR 18-DEC-1997; 97US-0068053.
XX PR 18-DEC-1997; 97US-0068054.
XX PR 18-DEC-1997; 97US-0068057.

```



PR 18-DEC-1997: 97US-0068064.  
PR 18-DEC-1997: 97US-0070923.  
PR 19-DEC-1997: 97US-0068169.  
PR 19-DEC-1997: 97US-0068365.  
PR 19-DEC-1997: 97US-0068367.  
PR 19-DEC-1997: 97US-0068368.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Carter KC, Duan RD, Feng P, Ferrle AM, Florence C;  
PI Florence K, Greene JM, Janat F, Kyaw H, Moore PA;  
PI Ni J, Rosen CA, Ruben SM, Shi Y, Soppet DR, Wei Y;  
PI Yu G;  
XX  
DR WPI: 1999-418749/35.  
XX  
PT New isolated human genes encoding secreted polypeptides  
XX  
PS Disclosure: Page 433; 537pp; English.  
XX  
CC AAX97916 to AAX98029 represent 110 isolated human secreted protein  
CC genes AAX936224 to AAX96727 represent the secreted proteins encoded by  
CC the 110 human genes. The genes and their corresponding secreted  
CC polypeptides are useful for preventing, treating or ameliorating medical  
CC conditions, e.g. by protein or gene therapy. Also pathological conditions  
CC can be diagnosed by determining the amount of the new polypeptides in a  
CC sample or by determining the presence of mutations in the new genes.  
CC Specific uses are described for each of the 110 genes, based on which  
CC tissues they are most highly expressed in, and include developing  
CC products for the diagnosis or treatment of cancer, tumours, developmental  
CC abnormalities and foetal deficiencies, blood disorders, diseases of the  
CC immune system, autoimmune diseases, inflammation, allergies, Alzheimer's  
CC and cognitive disorders, schizophrenia, arthritis, asthma, psoriasis,  
CC sepsis, skin disorders, atherosclerosis, diabetes, cardiovascular  
CC disorders, kidney disorders, digestive/endocrine disorders, infections  
CC and AIDS. The polypeptides are also useful for identifying their binding  
CC partners. The sequences given in AAX97907 to AAX97915 and AAX936223 are  
CC used in the exemplification of the present invention.  
XX  
SQ Sequence 24 AA:  
Query Match 48.5%; Score 32; DB 20; Length 24;  
Best Local Similarity 53.8%; Pred. No. 47;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 KLVPFATELHERL 13  
DB 12 KVPASASQLEFVRL 24  
RESULT 26  
ABBA6989  
ID ABA6989 standard; Peptide: 8 AA.  
XX  
AC ABA6989;  
XX  
XX 30-JAN-2002 (first entry).  
XX  
DE Desmocollin-2 CAR cyclic peptide 71.  
XX  
XX Desmosomal cadherin; cell adhesion; CAR sequence; immunosuppressive;  
XX cytosolic; antiapoptotic; wound healing; reduce scar tissue;  
XX skin graft; organ implant; autoimmune blistering disorder; cancer;  
XX apoptosis; cyclic.  
XX  
OS Synthetic.  
XX  
XX WO200172956-A2.  
XX  
XX 04-OCT-2001.  
XX  
XX 27-MAR-2001; 2001WO-IB01400.  
XX

PR 27-MAR-2000; 2000US-0535852.  
XX  
PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
XX  
PI Blaschuk OM, Symonds JM, Gour BJ;  
XX  
XX WPI: 2002-025778/03.  
XX  
DR  
XX  
PT Modulating agents for inhibiting or enhancing desmosomal cadherin  
PT mediated cell adhesion, useful for facilitating wound healing and/or  
PT reducing scar tissue, treating cancer and inducing apoptosis.  
XX  
PS Claim 23; Page 111; 127pp; English.  
XX  
CC The invention relates to modulating agents for inhibiting or enhancing  
CC desmosomal cadherin mediated cell adhesion, comprising a modulating agent  
CC comprising a desmosomal cadherin cell adhesion recognition CAR sequence  
CC (ABA45341-ABA47262), a non-peptide mimetic of a desmosomal cadherin CAR  
CC sequence, a substance such as an antibody or antigen-binding fragment  
CC that specifically binds a desmosomal cadherin CAR sequence and/or a  
CC polynucleotide encoding a polypeptide that comprises a desmosomal  
CC cadherin CAR sequence or analogue. The modulating agents have  
CC immunosuppressive, cytostatic and antiapoptotic activity and are used to  
CC facilitate wound healing and/or reduce scar tissue, for enhancing  
CC adhesion of foreign tissue implants (e.g. skin graft or organ implant),  
CC treating an autoimmune blistering disorder and to treat cancer  
CC (e.g. carcinoma, leukaemia or melanoma) and induce apoptosis.  
XX  
SQ Sequence 8 AA:  
Query Match 47.0%; Score 31; DB 23; Length 8;  
Best Local Similarity 75.0%; Pred. No. 7.8e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KLVPFATE 8  
DB 1 KLIAFATE 8  
RESULT 27  
AAB11444  
ID AAB11444 standard; peptide: 16 AA.  
XX  
AC AAB11444;  
XX  
XX 23-FEB-2001 (first entry)  
XX  
DT Interferon-alpha2 derived peptide monomer I.  
XX  
DE  
XX  
XX Interferon-alpha2; peptide monomer; immunosuppressive; cytostatic;  
XX anti-HIV; antiinflammatory; antiarthritic; antiallergic; dermatologic;  
XX antiproliferative; treatment; immune system disease; cell proliferation;  
XX cancer; capillary cell leukemia; metastasizing renal cell carcinoma;  
XX Kaposi sarcoma; lymphoma; sarcoma; arthritis; allergic reaction;  
XX skin disease; psoriasis.  
XX  
XX Unidentified.  
XX  
XX  
XX WO200064936-A2.  
XX  
XX 02-NOV-2000.  
XX  
XX 19-APR-2000; 2000WO-DE01234.  
XX  
XX 27-APR-1999; 99DE-1019149.  
XX  
XX (WIES/) WIESER R J.  
XX  
XX Wieser RJ;  
XX  
XX WPI: 2001-024674/03.  
XX  
XX New synthetic homo- or heterodimer peptides which bind to the  
PT

PT interferon alpha-2 receptor, useful e.g. for treating cancer, immune  
 PT dysfunction or inflammation or as diagnostic agents -  
 XX  
 PS Claim 1, Page 22, 26pp; German.

This invention describes novel synthetic homo- or heterodimers (A), which bind to the interferon alpha2 receptor. The products of the invention have immunosuppressive, cytostatic, anti-HIV (human immunodeficiency virus), antiinflammatory, antitubercitic, antiallergic, dermatological and antiparasitic activity. (A) is used for the treatment of diseases of the immune system, diseases dependent on elevated cell proliferation, diseases associated with elevated interferon synthesis or elevated expression of interferon alpha2 receptors, infectious or inflammatory processes, especially cancer diseases particularly capillary cell leukemia, metastasizing renal cell carcinoma, acquired immunodeficiency syndrome (AIDS)-associated Kaposi sarcoma, lymphoma or sarcoma. Other disorders which can be treated include arthritic processes, allergic reactions and skin diseases such as psoriasis. (A), or the antibodies against (A), can be used for the diagnosis of diseases associated with modified, insufficient or excessive expression of the interferon alpha2 receptor or insufficient or excessive interferon alpha2 concentrations. (A) has reduced side-effects, a longer half-life and higher biological activity compared with prior art drugs having interferon alpha2-related activity.

Sequence 16 AA;

Query Match	47.08;	Score 31;	DB 22;	Length 16;
Best Local Similarity	50.08;	Pred. No. 47;		
Matches 6; Conservative	4;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY      2 LVPFATELHERL 13
        | : | | | : : |
Db      5 LKFEYTELYQQL 16
```

RESULT 28  
ABB03758  
ID ABB03758 standard; Protein: 43 AA.

AC	ABB03758;
XX	
DT	08-JAN-2002 (first entry)
..	

Human musculoskeletal system related polypeptide SEQ ID NO 1705.

KM Cytostatic; immunosuppressive; nontropic; neuroprotective; antiviral;  
KM antiallergic; hepatotropic; antibacterial; antiinflammatory; antitumor;  
KM vulnerary; anticonvulsant; antibacterial; antitumoral; antiparasitic;  
KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KM neurological disease; infection; human; secreted protein;  
KM musculoskeletal system.

OS Homo sapiens.

PN WO200155367-A1.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01338.

PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0200545.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0209467.  
PR 30-JUN-2000; 2000US-0214886.  
PR 30-JUL-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.

PR	07-JUL-2000	2000US-0216680
PR	11-JUL-2000	2000US-0217487
PR	11-JUL-2000	2000US-0217496
PR	14-JUL-2000	2000US-0218290
PR	26-JUL-2000	2000US-0220963
PR	26-JUL-2000	2000US-0220964
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225267
PR	14-AUG-2000	2000US-0225268
PR	14-AUG-2000	2000US-0225270
PR	14-AUG-2000	2000US-0225447
PR	14-AUG-2000	2000US-0225757
PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759
PR	18-AUG-2000	2000US-0226279
PR	22-AUG-2000	2000US-0226681
PR	22-AUG-2000	2000US-0226688
PR	22-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227009
PR	30-AUG-2000	2000US-0228924
PR	01-SEP-2000	2000US-0228287
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229309
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231242
PR	08-SEP-2000	2000US-0231243
PR	08-SEP-2000	2000US-0231244
PR	08-SEP-2000	2000US-0231413
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0230801
PR	08-SEP-2000	2000US-0230802
PR	12-SEP-2000	2000US-0231967
PR	14-SEP-2000	2000US-0233397
PR	14-SEP-2000	2000US-0233398
PR	14-SEP-2000	2000US-0233399
PR	14-SEP-2000	2000US-0233400
PR	14-SEP-2000	2000US-0233401
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234274
PR	21-SEP-2000	2000US-0234997
PR	25-SEP-2000	2000US-0234998
PR	25-SEP-2000	2000US-0235484
PR	27-SEP-2000	2000US-0235834
PR	27-SEP-2000	2000US-0235836
PR	29-SEP-2000	2000US-0236327
PR	29-SEP-2000	2000US-0236367
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236369
PR	02-OCT-2000	2000US-0236802
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	13-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241121
PR	20-OCT-2000	2000US-0241175
PR	20-OCT-2000	2000US-0241176
PR	20-OCT-2000	2000US-0241808

PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 01-DEC-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 05-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI: 2001-451937/48.  
 DR N-PSDB; AAL35340.  
 XX  
 XX Isolated polypeptide for treating, preventing and/or prognosing  
 PT disorders related to the musculoskeletal system including  
 PT musculoskeletal cancers and also for testing and detection e.g.  
 PT diagnosis -  
 XX  
 XX Claim 11: SEQ ID NO 1705; 781bp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel genes (AAL34669-AAL37666) and proteins  
 CC (AB03087-AB04109) associated with the musculoskeletal system useful  
 CC for preventing, treating or ameliorating medical conditions e.g. by  
 CC protein or gene therapy. The genes are isolated from a range of human  
 CC tissues disclosed in the specification. The nucleic acids, proteins,  
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment  
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and

CC other cancers of the adrenal gland, bone, bone marrow, breast,  
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
 CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound  
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
 CC and (f) infectious diseases such as viral, bacterial, fungal and  
 CC parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SO Sequence 43 AA:  
 QY Query Match 47.0%; Score 31; DB 22; Length 43;  
 Db Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 VPFALEHER 12  
 Db 10 VPFLEHHSK 19  
 III I I I  
 RESULT 29  
 ID AAY30886  
 ID AAY30886 standard; Protein: 47 AA.  
 AC AAY30886;  
 AC  
 DT 12-OCT-1999 (first entry)  
 DT  
 XX  
 DE human secreted protein fragment encoded from gene 15.  
 XX  
 KW Secreted protein; prevention; treatment; protein therapy; gene therapy;  
 KW diagnosis; cancer; tumour; neurodegenerative disorder; blood disorder;  
 KW developmental abnormality; fetal deficiency; leukemia; autoimmune; acne;  
 KW hepatic disease; renal disease; lymphoma; inflammation; allergy; asthma;  
 KW Alzheimer's disease; cognitive disorder; schizophrenia; obesity; sepsis;  
 KW osteoporosis; arthritis; infection; AIDS; connective tissue disorder;  
 KW transplant rejection; diabetes; psoriasis; cardiovascular disorder;  
 KW reproductive disorder; food additive; food preservative; human; primer;  
 KW early promoter; GAS; gamma activation element.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO9440100-A1.  
 PN  
 XX 12-AUG-1999.  
 PD  
 XX  
 PF 04-FEB-1999; 99WO-US02293.  
 PF  
 XX  
 PR 09-FEB-1998; 98US-0074341.  
 PR 09-FEB-1998; 98US-0074037.  
 PR 09-FEB-1998; 98US-0074118.  
 PR 09-FEB-1998; 98US-0074141.  
 PR 09-FEB-1998; 98US-0074157.  
 PR  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Kraw H, Lafleur DW, Moore PA, Rosen CA, Ruben SM;  
 PI Shi Y, Wei Y;  
 XX  
 DR WPI: 1999-479426/40.  
 DR N-PSDB; AA200816.  
 XX  
 XX New isolated human genes potentially useful for, e.g. developmental  
 PT abnormalities and fetal deficiencies  
 PT  
 XX Disclosure; Page 248-249; 263pp; English.  
 PS  
 XX This invention describes novel isolated human genes and the secreted  
 CC proteins they encode. The polynucleotides and their corresponding

CC secreted polypeptides are useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. Also pathological  
CC conditions can be diagnosed by determining the amount of the new  
CC polypeptides in a sample or by determining the presence of mutations in  
CC the new polynucleotides. Specific uses are described for the  
CC polynucleotides of the invention based on which tissues they are most  
CC highly expressed in, and include developing products for the diagnosis or  
CC treatment of cancer, tumours, neurodegenerative disorders, developmental  
CC abnormalities and fetal deficiencies, blood disorders, leukemias,  
CC diseases of the immune system, autoimmune diseases, hepatic and renal  
CC diseases, lymphomas, inflammation, allergies, Alzheimer's and cognitive  
CC disorders, schizophrenia, obesity, osteoporosis, arthritis, infections,  
CC AIDS, connective tissue disorders, transplant rejection, diabetes,  
CC asthma, sepsis, acne, psoriasis, cardiovascular disorders, and  
CC reproductive disorders. The polypeptides or polynucleotides can also be  
CC used as food additives or preservatives. The polypeptide are also useful  
CC for identifying their binding partners. This sequence represents a  
CC secreted protein fragment described in the invention.

SQ Sequence 47 AA;

Query Match 47.0%; Score 31; DB 20; Length 47;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ELHERL 13

Db 8 ELHERL 13

RESULT 30

ABG01817 ID ABG01817 standard; Protein; 49 AA.

AC ABG01817;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #1808.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Dermanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR N-PSDB; AAS66004.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

PS Claim 20; SEQ ID NO 32176; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 49 AA;

Query Match 47.0%; Score 31; DB 22; Length 49;

Best Local Similarity 66.7%; Pred. No. 1.6e+02;

Matches 8; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 4 PFAT--ELHERL 13

Db 36 PIATASLHERL 47

Search completed: July 1, 2003, 17:56:21  
Job time : 36 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 17:55:45 ; Search time 14 Seconds  
(without alignments)  
27.321 Million cell updates/sec

Title: US-09-993-366-1  
Perfect score: 66  
Sequence: 1 KLVFFATELHERL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 160334

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 75 summaries

Database :

1: /cgn2\_6/ptodata/1/laa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/laa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/laa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/laa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/laa/PTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/laa/backfilest.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	48.5	15	4	US-08-766-596A-55
2	32	48.5	20	4	US-08-861-338-7
3	31	47.0	9	4	US-08-766-596A-52
4	30	45.5	35	4	US-09-149-476-337
5	30	45.5	43	4	US-09-149-476-638
6	29	43.9	20	1	US-07-961-724C-7
7	29	43.9	23	1	US-07-961-724C-6
8	29	43.9	34	4	US-08-469-260A-406
9	28	42.4	18	1	US-08-362-453-2
10	28	42.4	23	1	US-08-485-718-3
11	28	42.4	23	2	US-08-484-530-3
12	28	42.4	23	2	US-08-827-618A-3
13	28	42.4	23	3	US-08-483-952A-3
14	28	42.4	23	3	US-08-476-501-3
15	28	42.4	27	1	US-08-276-936A-24
16	28	42.4	27	5	PCT-US95-04488-24
17	28	42.4	31	2	US-08-637-759B-500
18	28	42.4	31	3	US-08-871-355A-500
19	28	42.4	31	4	US-09-201-945-500
20	28	42.4	34	1	US-08-438-753B-40
21	28	42.4	34	1	US-08-443-883A-40
22	28	42.4	34	2	US-08-631-328-40
23	28	42.4	34	2	US-08-455-524B-40
24	28	42.4	34	2	US-08-455-021B-40
25	28	42.4	34	4	US-09-045-467-40
26	28	42.4	34	4	US-08-934-915-40
27	27	40.9	20	2	US-08-934-915-40

28	27	40.9	20	2	US-08-934-915-170	Sequence 170, App
29	27	40.9	25	4	US-09-314-268-4	Sequence 4, Appl
30	27	40.9	30	2	US-08-934-915-42	Sequence 42, Appl
31	27	40.9	32	6	5182210-22	Patent No. 5182210
32	27	40.9	35	4	US-09-314-268-117	Sequence 117, App
33	27	40.9	50	6	5182210-16	Patent No. 5182210
34	26	39.4	5	4	US-08-766-596A-32	Sequence 32, Appl
35	26	39.4	12	1	US-08-190-788A-159	Sequence 159, App
36	26	39.4	12	1	US-08-383-474B-162	Sequence 162, App
37	26	39.4	12	1	US-08-465-538A-159	Sequence 159, App
38	26	39.4	12	1	US-08-464-538B-159	Sequence 159, App
39	26	39.4	12	2	US-08-463-076E-211	Sequence 211, App
40	26	39.4	14	1	US-08-408-604A-21	Sequence 21, Appl
41	26	39.4	14	5	PCT-US93-09626-21	Sequence 21, Appl
42	26	39.4	15	4	US-09-561-490E-18	Sequence 18, Appl
43	26	39.4	19	4	US-09-292-225-11	Sequence 11, Appl
44	26	39.4	20	4	US-09-292-225-3	Sequence 3, Appl
45	26	39.4	26	2	US-08-563-892A-17	Sequence 17, Appl
46	26	39.4	27	2	US-08-563-892A-16	Sequence 16, Appl
47	25	37.9	8	6	5310729-42	Patent No. 5310729
48	25	37.9	10	1	US-08-131-057A-6	Sequence 6, Appl
49	25	37.9	15	4	US-09-230-405-13	Sequence 13, Appl
50	25	37.9	16	2	US-08-934-222-26	Sequence 26, Appl
51	25	37.9	16	2	US-08-933-402-26	Sequence 26, Appl
52	25	37.9	16	2	US-09-207-621-26	Sequence 26, Appl
53	25	37.9	16	2	US-08-532-818-26	Sequence 26, Appl
54	25	37.9	16	2	US-08-934-223-26	Sequence 26, Appl
55	25	37.9	16	3	US-09-231-797-26	Sequence 26, Appl
56	25	37.9	16	3	US-08-934-224-26	Sequence 26, Appl
57	25	37.9	16	3	US-08-933-843-26	Sequence 26, Appl
58	25	37.9	16	4	US-09-413-492-26	Sequence 26, Appl
59	25	37.9	20	3	US-08-335-865J-19	Sequence 19, Appl
60	25	37.9	25	1	US-08-345-527-5	Sequence 5, Appl
61	25	37.9	25	3	US-08-887-769-5	Sequence 5, Appl
62	25	37.9	43	4	US-08-905-223-293	Sequence 293, App
63	25	37.9	46	4	US-09-461-697-50	Sequence 50, Appl
64	25	37.9	47	2	US-08-766-858A-12	Sequence 12, Appl
65	24.5	37.1	24	1	US-07-908-317-9	Sequence 9, Appl
66	24.5	37.1	24	5	PCT-US93-06171-9	Sequence 9, Appl
67	24.5	37.1	38	1	US-08-176-500-84	Sequence 84, Appl
68	24.5	37.1	38	1	US-08-471-052A-84	Sequence 84, Appl
69	24.5	37.1	38	1	US-08-189-331-84	Sequence 84, Appl
70	24.5	37.1	38	2	US-08-471-939-84	Sequence 84, Appl
71	24.5	37.1	38	2	US-08-471-800-84	Sequence 84, Appl
72	24.5	37.1	38	2	US-08-471-068-84	Sequence 84, Appl
73	24.5	37.1	40	2	US-08-194-961E-25	Sequence 25, Appl
74	24.5	37.1	47	1	US-08-582-257-14	Sequence 14, Appl
75	24.5	37.1	47	2	US-08-582-298-14	Sequence 14, Appl

#### ALIGNMENTS

RESULT 1  
US-08-766-596A-55  
Sequence 55, Application US/08766596A  
Patent No. 6462171  
GENERAL INFORMATION:  
APPLICANT: SOTO-JARA, Claudio  
APPLICANT: BAUMANN, Marc  
APPLICANT: FRANGIONE, Blas  
TITLE OF INVENTION: PEPTIDES AND PHARMACEUTICAL  
COMPOSITIONS THEREOF FOR TREATMENT OF DISORDERS OR DISEASES  
ASSOCIATED WITH PROTEIN FOLDING INTO AMYLOID OR AMYLOID-LIK  
TITLE OF INVENTION: ASSOCIATED WITH PROTEIN FOLDING INTO AMYLOID OR AMYLOID-LIK  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 400  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/766,596A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/630,645  
FILING DATE: 10-APR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/478,326  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: YUN, Allen C.  
REGISTRATION NUMBER: 37,971  
REFERENCE/DOCKET NUMBER: SOTO-JARA-1A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-766-596A-55

Query Match 48.5%; Score 32; DB 4; Length 15;  
Best Local Similarity 66.7%; Pred. No. 7.8;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVPEATEL 9  
Db 5 KLVPEADV 13

RESULT 2  
US-08-861-338-7  
Sequence 7, Application US/08861338  
Patent No. 6174993  
GENERAL INFORMATION:  
APPLICANT: Ben-Sasson, Shmuel A.  
TITLE OF INVENTION: SHORT PEPTIDES WHICH SELECTIVELY  
MODULATE THE ACTIVITY OF SERINE/THREONINE KINASES  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Milltia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/861,338  
FILING DATE: 21-MAY-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CMCC-590  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 861-6240  
TELEFAX: (781) 861-9540  
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-861-338-7

Query Match 48.5%; Score 32; DB 4; Length 20;  
Best Local Similarity 54.5%; Pred. No. 11;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 VPFAELHERL 13  
Db 8 LPFYQDHERL 18

RESULT 3  
US-08-766-596A-52  
Sequence 52, Application US/08766596A  
Patent No. 6462171  
GENERAL INFORMATION:  
APPLICANT: SOTO-JARA, Claudio  
APPLICANT: BAUMANN, Marc  
APPLICANT: FRANGIONE, Blas  
TITLE OF INVENTION: PEPTIDES AND PHARMACEUTICAL  
COMPOSITIONS THEREOF FOR TREATMENT OF DISORDERS OR DISEASES  
ASSOCIATED WITH PROTEIN FOLDING INTO AMYLOID OR AMYLOID-LIK  
TITLE OF INVENTION: DEPOSITS  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 400  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/766,596A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/630,645  
FILING DATE: 10-APR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/478,326  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: YUN, Allen C.  
REGISTRATION NUMBER: 37,971  
REFERENCE/DOCKET NUMBER: SOTO-JARA-1A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-766-596A-52

Query Match 47.0%; Score 31; DB 4; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.9e+05;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KLVPEATE 8

Db 2 KLVFPAED 9

RESULT 4

US-09-149-476-337

Sequence 337, Application US/09149476

Patent No. 6420526

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002P1

CURRENT FILING DATE: 1998-09-08

EARLIER APPLICATION NUMBER: PCT/US98/04493

EARLIER FILING DATE: 1998-03-06

EARLIER APPLICATION NUMBER: 60/040,162

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,333

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/038,621

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,626

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,334

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,336

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,163

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/047,600

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,615

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,597

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,502

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,633

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,583

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,617

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,618

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,503

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,592

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,581

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,584

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,500

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,587

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,492

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,598

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,613

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,582

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,596

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,612

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,632

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,601

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,601

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,601

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,601

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,601

```
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match      45.58; Score 30; DB 4; Length 35;
Best Local Similarity 41.7%; Pred. No. 47;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
OY      2  LVPATLHERL 13
      :  :  :  :  :  :
Db      18  ILPVCALHERL 29
```

```
RESULT 5
US-09-149-476-638
; Sequence 638, Application US/09149476
```

```
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
```



EARLIER APPLICATION NUMBER: 60/043,569  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,311  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588

EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 45.5%; Score 30; DB 4; Length 43;  
Best Local Similarity 41.7%; Pred. No. 58;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 LVPEATELHERL 13  
Db 30 ILPVCAHLHEEL 41

RESULT 6  
US-07-961-724C-7

Sequence 7, Application US/07961724C  
Patent No. 5541078

GENERAL INFORMATION:

APPLICANT: FACON, BRIGITTE  
APPLICANT: CHAMEKH, MUSTAPHA  
APPLICANT: DISSOUS, COLETTE  
APPLICANT: CAPRON, ANDRE  
APPLICANT: TARTAR, ANDRE

APPLICANT: GRAS-MASSÉ, HELENE  
TITLE OF INVENTION: IMMUNOGENIC PEPTIDE SEQUENCE OF  
TITLE OF INVENTION: ECHINOCOCCUS GRANULOSUS, DNA SEQUENCE CODING FOR THIS  
TITLE OF INVENTION: PEPTIDE SEQUENCE AND DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: APPLICATIONS  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/961,724C  
FILING DATE: 10-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 90/08900  
FILING DATE: 12-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, No. 5541078man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 660-065-0X PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-07-961-724C-7

Query Match 43.9%; Score 29; DB 1; Length 20;  
Best Local Similarity 30.8%; Pred. No. 39;  
Matches 4; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVFATELHERL 13  
DB 3 KMLTEASDVHQR 15

RESULT 7  
US-07-961-724C-6  
Sequence 6, Application US/07961724C  
Patent No. 5541078  
GENERAL INFORMATION:  
APPLICANT: FACON, BRIGITTE  
APPLICANT: CHAMEKH, MUSTAPHA  
APPLICANT: DISSOUS, COLETTE  
APPLICANT: CAPRON, ANDRE  
APPLICANT: TARTAR, ANDRE  
APPLICANT: GRAS-MASSÉ, HELENE  
TITLE OF INVENTION: IMMUNOGENIC PEPTIDE SEQUENCE OF  
TITLE OF INVENTION: ECHINOCOCCUS GRANULOSUS, DNA SEQUENCE CODING FOR THIS  
TITLE OF INVENTION: PEPTIDE SEQUENCE AND DIAGNOSTIC AND THERAPEUTIC  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,

ADDRESSEE: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/961,724C  
FILING DATE: 10-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 90/08900  
FILING DATE: 12-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, No. 5541078man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 660-065-0X PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-07-961-724C-6

Query Match 43.9%; Score 29; DB 1; Length 23;  
Best Local Similarity 30.8%; Pred. No. 45;  
Matches 4; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVFATELHERL 13  
DB 8 KMLTEASDVHQR 20

RESULT 8  
US-08-469-260A-406  
Sequence 406, Application US/08469260A  
Patent No. 6451578  
GENERAL INFORMATION:  
APPLICANT: JOHN N. SIMONS  
APPLICANT: TAM J. PILOT-MATIAS  
APPLICANT: GEORGE J. DAMSON  
APPLICANT: GEORGE G. SCHLAUDER  
APPLICANT: SURESH M. DESAI  
APPLICANT: THOMAS P. LEARY  
APPLICANT: ANTHONY SCOTT MUEHRHOF  
APPLICANT: JAMES C. ERKER  
APPLICANT: SHERI L. BUTIK  
APPLICANT: ISA K. MUSHAMMAR  
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS  
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE  
NUMBER OF SEQUENCES: 716  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D  
STREET: 100 ABBOTT PARK ROAD  
CITY: ABBOTT PARK  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,260A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/424,550  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: POREMBSKI, PRISCILLA E.  
REGISTRATION NUMBER: 33,207  
REFERENCE/DOCKET NUMBER: 5527.PC.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708-937-6365  
TELEFAX: 708-938-2623  
INFORMATION FOR SEQ ID NO: 406:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-469-260A-406

Query Match 43.9%; Score 29; DB 4; Length 34;  
Best Local Similarity 45.5%; Pred. No. 70;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LVFPEATELHER 12  
|:|:|:|:|  
Db 13 LLPMANDAHVR 23

RESULT 9  
US-08-362-453-2  
Sequence 2, Application US/08362453  
Patent No. 5684129  
GENERAL INFORMATION:  
APPLICANT: FISH, Eleanor N.  
TITLE OF INVENTION: INTERFERON RECEPTOR BINDING PEPTIDES  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikaïdo, Marmelstein, Murray & Oram  
STREET: 655 Fifteenth Street N.W. Suite 330  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/362,453  
FILING DATE: 06-JAN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/909,739  
APPLICATION NUMBER: FILING DATE: 07-JUL-1992  
APPLICATION NUMBER: US 07/980,525  
APPLICATION NUMBER: FILING DATE: 20-NOV-1992  
APPLICATION NUMBER: PCT/CA93/00279  
APPLICATION NUMBER: FILING DATE: 06-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kilts, Monica Chin  
REGISTRATION NUMBER: 36,105  
REFERENCE/DOCKET NUMBER: P638-4017  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-5000  
TELEFAX: (202) 638-4810

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-362-453-2

Query Match 42.4%; Score 28; DB 1; Length 18;  
Best Local Similarity 46.2%; Pred. No. 53;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LVFPEATELHERL 13  
|:|:|:|:|  
Db 5 LEKFTYELLYOOL 16

RESULT 10  
US-08-485-718-3  
Sequence 3, Application US/08485718  
Patent No. 5705626  
GENERAL INFORMATION:  
APPLICANT: Tobin, Allan J.  
APPLICANT: Erlander, Mark G.  
APPLICANT: Kaufman, Daniel L.  
TITLE OF INVENTION: Cloned Glutamic Acid Decarboxylase  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,718  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Rowland, Bertram I.  
REGISTRATION NUMBER: 20015  
REFERENCE/DOCKET NUMBER: A60780-6/BIR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-718-3

Query Match 42.4%; Score 28; DB 1; Length 23;  
Best Local Similarity 46.2%; Pred. No. 70;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVFPEATELHERL 13  
|:|:|:|:|  
Db 8 KILPEVKEKHEFL 20

RESULT 11  
US-08-484-530-3  
Sequence 3, Application US/08484530  
Patent No. 5846740  
GENERAL INFORMATION:

APPLICANT: Tobin, Allan J  
APPLICANT: Erlander, Mark G  
TITLE OF INVENTION: Cloned Glutamic Acid Decarboxylase  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,530  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-530-3

Query Match 42.4%; Score 28; DB 2; Length 23;  
Best Local Similarity 46.2%; Pred. No. 70;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13  
Db 8 KILPEVKEKHEFL 20

RESULT 12  
US-08-827-618A-3  
Sequence 3, Application US/08827618A  
Patent No. 5998366  
GENERAL INFORMATION:  
APPLICANT: Tobin, Allan J  
APPLICANT: Erlander, Mark G  
TITLE OF INVENTION: Cloned Glutamic Acid Decarboxylase  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/827,618A  
FILING DATE: 09-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/485,725  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/716,909  
FILING DATE: 18-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/586,536  
FILING DATE: 21-SEP-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-827-618A-3

Query Match 42.4%; Score 28; DB 2; Length 23;  
Best Local Similarity 46.2%; Pred. No. 70;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13  
Db 8 KILPEVKEKHEFL 20

RESULT 13  
US-08-483-952A-3  
Sequence 3, Application US/08483952A  
Patent No. 6011139  
GENERAL INFORMATION:  
APPLICANT: Tobin, Allan J  
APPLICANT: Erlander, Mark G  
TITLE OF INVENTION: Cloned Glutamic Acid Decarboxylase  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,952A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/123,859  
FILING DATE: 17-SEP-1993  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/716,909  
FILING DATE: 18-JUN-1991  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/586,536  
FILING DATE: 21-SEP-1990  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.

REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-60780-7/RFT/MTK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-483-952A-3

Query Match 42.4% Score 28; DB 3; Length 23;  
Best Local Similarity 46.2% Pred. No. 70;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVFATLHRL 13  
|:|:| | | | |  
Db 8 KILPEVKEKHEFL 20

RESULT 14  
US-08-476-501-3  
Sequence 3, Application US/08476501  
Patent No. 6455267  
GENERAL INFORMATION:  
APPLICANT: TOBIN, Allan J.  
APPLICANT: ERLANDER, Mark G.  
APPLICANT: KAUFMAN, Daniel L.  
TITLE OF INVENTION: Cloned Glutamic Acid Decarboxylase  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,501  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rowland, Bertram I.  
REGISTRATION NUMBER: 20015  
REFERENCE/DOCKET NUMBER: A60780-10/BIR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-476-501-3

Query Match 42.4% Score 28; DB 4; Length 23;  
Best Local Similarity 46.2% Pred. No. 70;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVFATLHRL 13  
|:|:| | | | |  
Db 8 KILPEVKEKHEFL 20

RESULT 15  
US-08-276-936A-24  
Sequence 24, Application US/08276936A  
Patent No. 5612194  
GENERAL INFORMATION:  
APPLICANT: Harvey Rubin, Barry Cooperman, No. 5612194man Schecter,  
APPLICANT: Michael Plotkin, Zhi Wang  
TITLE OF INVENTION: Methods of Producing Effective  
TITLE OF INVENTION: Recombinant Serine Protease Inhibitors and Uses of These  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Jane Massey Licata, Esq.  
STREET: 210 Lake Drive East, Suite 201  
CITY: Cherry Hill  
STATE: NJ  
COUNTRY: USA  
ZIP: 08002

COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE  
COMPUTER: IBM 486  
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/276,936A  
FILING DATE: July 19, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 221,078  
FILING DATE: March 31, 1994  
APPLICATION NUMBER: 221,171  
FILING DATE: March 31, 1994  
APPLICATION NUMBER: 005,908  
FILING DATE: January 15, 1993  
APPLICATION NUMBER: 735,335  
FILING DATE: July 24, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane Massey Licata  
REGISTRATION NUMBER: 32,257  
REFERENCE/DOCKET NUMBER: PENN-0027  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (609) 779-2400  
TELEFAX: (609) 779-8488  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-276-936A-24

Query Match 42.4% Score 28; DB 1; Length 27;  
Best Local Similarity 33.3% Pred. No. 83;  
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVFATLHRL 12  
|:|:| | | | |  
Db 10 KIMPLSTOVRTR 21

RESULT 16  
US-08-722-268-24  
Sequence 24, Application US/08722268  
Patent No. 5827662  
GENERAL INFORMATION:  
APPLICANT: Harvey Rubin, Barry Cooperman, No. 5827662man Schecter,  
APPLICANT: Michael Plotkin, Zhi Wang  
TITLE OF INVENTION: Methods of Producing Effective  
TITLE OF INVENTION: Recombinant Serine Protease Inhibitors and Uses of These  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Jane Massey Licata, Esq.

STREET: 210 Lake Drive East, Suite 201  
CITY: Cherry Hill  
STATE: NJ  
COUNTRY: USA  
ZIP: 08002  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE  
COMPUTER: IBM 486  
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/722,268  
FILING DATE: December 18, 1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 276,936  
FILING DATE: July 19, 1994  
APPLICATION NUMBER: 221,078  
FILING DATE: March 31, 1994  
APPLICATION NUMBER: 221,171  
FILING DATE: March 31, 1994  
APPLICATION NUMBER: 005,908  
FILING DATE: January 15, 1993  
APPLICATION NUMBER: 735,335  
FILING DATE: July 24, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane Massey Licata  
REGISTRATION NUMBER: 32,257  
REFERENCE/DOCKET NUMBER: PENN-0107  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (609) 779-2400  
TELEFAX: (609) 779-8488  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-722-268-24

Query Match 42.4%; Score 28; DB 2; Length 27;  
Best Local Similarity 33.3%; Pred. No. 83;  
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPAETELHER 12  
|:|:|:|:|  
Db 10 KIMPLSTOETRR 21

RESULT 17  
PCT-US95-04488-24  
Sequence 24, Application PC/TUS9504488  
GENERAL INFORMATION:  
APPLICANT: Harvey Rubin, Barry Cooperman, Norman Schecter,  
APPLICANT: Michael Plotkin, Zhi Wang  
TITLE OF INVENTION: Methods of Producing Effective  
TITLE OF INVENTION: Recombinant Serine Protease Inhibitors and Uses of These  
NUMBER OF INVENTIONS: 34  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Jane Massey Licata, Esq.  
STREET: 210 Lake Drive East, Suite 201  
CITY: Cherry Hill  
STATE: NJ  
COUNTRY: USA  
ZIP: 08002  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE  
COMPUTER: IBM 486  
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04488  
FILING DATE: Herewith

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 276,936  
FILING DATE: July 19, 1994  
APPLICATION NUMBER: 221,078  
FILING DATE: March 31, 1994  
APPLICATION NUMBER: 221,171  
FILING DATE: March 31, 1994  
APPLICATION NUMBER: 005,908  
FILING DATE: January 15, 1993  
APPLICATION NUMBER: 735,335  
FILING DATE: July 24, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane Massey Licata  
REGISTRATION NUMBER: 32,257  
REFERENCE/DOCKET NUMBER: PENN-0107  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (609) 779-2400  
TELEFAX: (609) 779-8488  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27  
TYPE: Amino Acid  
TOPOLOGY: Linear  
PCT-US95-04488-24

Query Match 42.4%; Score 28; DB 5; Length 27;  
Best Local Similarity 33.3%; Pred. No. 83;  
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPAETELHER 12  
|:|:|:|:|  
Db 10 KIMPLSTOETRR 21

RESULT 18  
US-08-637-759B-500  
Sequence 500, Application US/08637759B  
Patent No. 5876931  
GENERAL INFORMATION:  
APPLICANT: David William Holden  
TITLE OF INVENTION: Identification of Genes  
NUMBER OF SEQUENCES: 501  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/637,759B  
FILING DATE: 03-MAY-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB95/02875  
FILING DATE: 11-DEC-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: RPMS 101  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 873-8794  
TELEFAX: (404) 873-8795  
INFORMATION FOR SEQ ID NO: 500:  
SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-637-759B-500

Query Match 42.4%; Score 28; DB 2; Length 31;  
Best Local Similarity 60.0%; Pred. No. 97;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 PFATLHERL 13  
11 1:111  
DB 12 PFILLYERL 21

RESULT 19  
US-08-871-355A-500  
; Sequence 500, Application US/08871355A  
; Patent No. 6015569  
; GENERAL INFORMATION:  
; APPLICANT: David William Holden  
; TITLE OF INVENTION: Identification of Genes  
; NUMBER OF SEQUENCES: 501  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center  
; STREET: 1201 West Peachtree Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/871.355A  
; FILING DATE: 09-JUN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/02875  
; FILING DATE: 11-DEC-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: RPLMS 101 CON  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 873-8794  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 500:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
US-08-871-355A-500

Query Match 42.4%; Score 28; DB 3; Length 31;  
Best Local Similarity 60.0%; Pred. No. 97;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 PFATLHERL 13  
11 1:111  
DB 12 PFILLYERL 21

RESULT 20

US-09-201-945-500  
; Sequence 500, Application US/09201945  
; Patent No. 6342215  
; GENERAL INFORMATION:  
; APPLICANT: David William Holden  
; TITLE OF INVENTION: Identification of Genes  
; NUMBER OF SEQUENCES: 501  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center  
; STREET: 1201 West Peachtree Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/201.945  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/637,759  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: RPLMS 101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 873-8794  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 500:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
US-09-201-945-500

Query Match 42.4%; Score 28; DB 4; Length 31;  
Best Local Similarity 60.0%; Pred. No. 97;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 PFATLHERL 13  
11 1:111  
DB 12 PFILLYERL 21

RESULT 21  
US-08-438-753B-40  
; Sequence 40, Application US/08438753B  
; Patent No. 5705363  
; GENERAL INFORMATION:  
; APPLICANT: Imakawa, Kazuhito  
; TITLE OF INVENTION: Interferon Tau Compositions and  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Denlinger & Associates  
; STREET: 350 Cambridge Ave., Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/438,753B
FILING DATE: 10-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/139,891
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,741
FILING DATE: 09-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/318,050
FILING DATE: 02-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/969,890
FILING DATE: 30-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 5600-0001.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-324-0880
TELEFAX: 415-324-0960
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
INDIVIDUAL ISOLATE: Amino acid sequence of fragment
US-08-438-753B-40
INDIVIDUAL ISOLATE: 139-172 of SEQ ID NO:33

Query Match
Best Local Similarity 42.4%; Score 28; DB 1; Length 34;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 ATELHERL 13
DB 17 STSLHRL 24

RESULT 22
US-08-443-883A-40
; Sequence 40, Application US/08443883A
; Patent No. 5738845
; GENERAL INFORMATION:
; APPLICANT: Bazer, Fuller W.
; APPLICANT: Johnson, Howard M.
; APPLICANT: Pontzer, Carol H.
; APPLICANT: Ott, Troy L.
; APPLICANT: Van Heeke, Gino
; APPLICANT: Imakawa, Kazuhiko
; TITLE OF INVENTION: Interferon Tau Compositions and
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/443,883A
```

```
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/139,891
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,741
FILING DATE: 09-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/318,050
FILING DATE: 02-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/969,890
FILING DATE: 30-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 5600-0001.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-324-0880
TELEFAX: 415-324-0960
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
INDIVIDUAL ISOLATE: Amino acid sequence of fragment
US-08-443-883A-40
INDIVIDUAL ISOLATE: 139-172 of SEQ ID NO:33

Query Match
Best Local Similarity 42.4%; Score 28; DB 1; Length 34;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 ATELHERL 13
DB 17 STSLHRL 24

RESULT 23
US-08-631-328-40
; Sequence 40, Application US/08631328
; Patent No. 5939286
; GENERAL INFORMATION:
; APPLICANT: Johnson, Howard M.
; APPLICANT: Pontzer, Carol H.
; APPLICANT: Subramaniam, Prem S.
; TITLE OF INVENTION: Hybrid Interferon Compositions and
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/631,328
; FILING DATE: 12-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/438,753
; FILING DATE: 10-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
```



REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 5600-0001.34  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-324-0880  
TELEFAX: 415-324-0960  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Amino acid sequence of fragment  
INDIVIDUAL ISOLATE: 139-172 of SEQ ID NO:33  
US-08-631-328-40

Query Match 42.4%; Score 28; DB 2; Length 34;  
Best Local Similarity 62.5%; Pred. NO. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 ATELHERL 13  
:111:11  
DB 17 STSLHKRL 24

RESULT 24  
US-08-455-524B-40  
Sequence 40, Application US/08455524B  
GENERAL INFORMATION:  
APPLICANT: Bazer, Fuller W.  
APPLICANT: Johnson, Howard M.  
APPLICANT: Pontzer, Carol H.  
APPLICANT: Ott, Troy L.  
APPLICANT: Van Heeke, Gino  
TITLE OF INVENTION: Interferon Tau Compositions and  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Ave., Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,524B  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/438,753  
FILING DATE: 10-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/139,891  
FILING DATE: 19-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847,741  
FILING DATE: 09-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/318,050  
FILING DATE: 02-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/969,890  
FILING DATE: 30-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 5600-0001.32

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-324-0880  
TELEFAX: 415-324-0960  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Amino acid sequence of fragment  
INDIVIDUAL ISOLATE: 139-172 of SEQ ID NO:33  
US-08-455-524B-40

Query Match 42.4%; Score 28; DB 2; Length 34;  
Best Local Similarity 62.5%; Pred. NO. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 ATELHERL 13  
:111:11  
DB 17 STSLHKRL 24

RESULT 25  
US-08-455-021B-40  
Sequence 40, Application US/08455021B  
GENERAL INFORMATION:  
APPLICANT: Bazer, Fuller W.  
APPLICANT: Johnson, Howard M.  
APPLICANT: Pontzer, Carol H.  
APPLICANT: Ott, Troy L.  
APPLICANT: Van Heeke, Gino  
APPLICANT: Imakawa, Kazuhito  
TITLE OF INVENTION: Interferon Tau Compositions and  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Ave., Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,021B  
FILING DATE: 31-MAY-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/139,891  
FILING DATE: 19-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847,741  
FILING DATE: 09-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/318,050  
FILING DATE: 02-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/969,890  
FILING DATE: 30-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 5600-0001.31  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-324-0880  
TELEFAX: 415-324-0960  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:

LENGTH: 34 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Amino acid sequence of fragment  
INDIVIDUAL ISOLATE: 139-172 of SEQ ID NO:33  
US-08-455-021B-40

Query Match 42.4%; Score 28; DB 2; Length 34;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 ATELHERL 13  
Db 17 STSLHRRL 24

RESULT 26  
US-09-045-467-40

; Sequence 40; Application US/09045467  
; GENERAL INFORMATION:

; APPLICANT: Johnson, Howard M.

; TITLE OF INVENTION: Interferon Tau Compositions and  
; Methods of Use

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Ave., Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/045,467  
; FILING DATE: 20-Mar-1998  
; CLASSIFICATION: <Unknown>

; PRIORITY INFORMATION:  
; APPLICATION NUMBER: US 08/455,021  
; FILING DATE: 31-May-1995  
; APPLICATION NUMBER: US 08/438,753  
; FILING DATE: 10-May-1995  
; APPLICATION NUMBER: US 08/139,891  
; FILING DATE: 19-Oct-1993  
; APPLICATION NUMBER: US 07/847,741  
; FILING DATE: 09-Mar-1992  
; APPLICATION NUMBER: US 07/318,050  
; FILING DATE: 02-Mar-1989  
; APPLICATION NUMBER: US 07/969,890  
; FILING DATE: 30-Oct-1992

; ATTORNEY/AGENT INFORMATION:  
; NAME: Dehlinger, Peter J.  
; REGISTRATION NUMBER: 28,006  
; REFERENCE/DOCKET NUMBER: 5600-0001.36  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-0880  
; TELEFAX: 650-324-0960

; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 34 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: Amino acid sequence of fragment  
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:

US-09-045-467-40

Query Match 42.4%; Score 28; DB 4; Length 34;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 ATELHERL 13  
Db 17 STSLHRRL 24

RESULT 27  
US-08-934-915-40

; Sequence 40; Application US/08934915  
; Patent No. 5932412  
; GENERAL INFORMATION:

; APPLICANT: DILLNER, JOAKIM  
; APPLICANT: DILLNER, LENA  
; APPLICANT: CHENG, HWEI-MING

; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
; TITLE OF INVENTION: USEFUL IN IMMUNOSSAY FOR  
; DIAGNOSTIC PURPOSES

; NUMBER OF SEQUENCES: 193  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MASON & ASSOCIATES, P.A.  
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
; CITY: CLEARWATER  
; STATE: FLORIDA  
; COUNTRY: U.S.A.

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: Windows 3.0  
; SOFTWARE: Microsoft Word 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/934,915  
; FILING DATE: 22-SEP-1997  
; CLASSIFICATION: 435  
; PRIORITY INFORMATION:  
; APPLICATION NUMBER: 07/949,836  
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:  
; NAME: LOUISE A. Foulch  
; REGISTRATION NUMBER: 37,133  
; REFERENCE/DOCKET NUMBER: 1946.6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 813-538-3800  
; TELEFAX: 813-538-3820  
; TELEX:

; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-934-915-40

Query Match 40.9%; Score 27; DB 2; Length 20;  
Best Local Similarity 50.0%; Pred. No. 92;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 PPEATLHERL 13  
Db 6 PWAPKKHRRL 15

RESULT 28  
US-08-934-915-170  
; Sequence 170; Application US/08934915  
; Patent No. 5932412  
; GENERAL INFORMATION:

APPLICANT: DILLNER, JOAKIM  
APPLICANT: DILLNER, LENA  
APPLICANT: CHENG, HWEI-MING  
TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR  
TITLE OF INVENTION: DIAGNOSTIC PURPOSES  
NUMBER OF SEQUENCES: 193  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MASON & ASSOCIATES, P. A.  
STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
CITY: CLEARWATER  
STATE: FLORIDA  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 3.0  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/934,915  
FILING DATE: 22-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/949,836  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: LOUISE A. FOUTCH  
REGISTRATION NUMBER: 37,133  
REFERENCE/DOCKET NUMBER: 1946.6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 813-538-3800  
TELEFAX: 813-538-3820  
TELEX:  
INFORMATION FOR SEQ ID NO: 170:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-934-915-170

Query Match 40.9%; Score 27; DB 2; Length 20;  
Best Local Similarity 50.0%; Pred. No. 92;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
1:1:11  
Db 6 PMAPKRRRL 15

RESULT 29  
US-09-314-268-4  
Sequence 4, Application US/09314268  
Patent No. 6346377  
GENERAL INFORMATION:  
APPLICANT: Doochar, John  
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA  
TITLE OF INVENTION: VIRUSES  
FILE REFERENCE: 3789/80902  
CURRENT APPLICATION NUMBER: US/09/314,268  
CURRENT FILING DATE: 1999-03-19  
EARLIER APPLICATION NUMBER: 09/314,268  
EARLIER FILING DATE: 1999-05-18  
NUMBER OF SEQ ID NOS: 179  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Human papillomavirus type 16  
US-09-314-268-4

Query Match 40.9%; Score 27; DB 4; Length 25;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
1:1:11  
Db 8 PMAPKRRRL 17

RESULT 30  
US-08-934-915-42  
Sequence 42, Application US/08934915  
Patent No. 5932412  
GENERAL INFORMATION:  
APPLICANT: DILLNER, JOAKIM  
APPLICANT: DILLNER, LENA  
APPLICANT: CHENG, HWEI-MING  
TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,  
TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR  
TITLE OF INVENTION: DIAGNOSTIC PURPOSES  
NUMBER OF SEQUENCES: 193  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MASON & ASSOCIATES, P. A.  
STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
CITY: CLEARWATER  
STATE: FLORIDA  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 3.0  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/934,915  
FILING DATE: 22-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/949,836  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: LOUISE A. FOUTCH  
REGISTRATION NUMBER: 37,133  
REFERENCE/DOCKET NUMBER: 1946.6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 813-538-3800  
TELEFAX: 813-538-3820  
TELEX:  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-934-915-42

Query Match 40.9%; Score 27; DB 2; Length 30;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
1:1:11  
Db 14 PMAPKRRRL 23

Search completed: July 1, 2003, 17:57:55  
Job time : 16 secs

•  
•  
•  
•

•

•

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 17:55:15 ; Search time 15 Seconds  
(Without alignments)  
83.316 Million cell updates/sec

Title: US-09-993-366-1  
Perfect score: 66  
Sequence: 1 KLVFPATELHERL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 11827

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database :  
1: PIR\_73:\*  
2: p1r1:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	50.0	29	2 F83870	hypothetical prote
2	29	43.9	37	2 S14101	apolipoprotein III
3	28	42.4	40	2 A61147	manganese peroxida
4	28	42.4	45	2 T52124	hypothetical prote
5	26	39.4	31	2 S15837	hypothetical prote
6	26	39.4	37	2 I48405	histone H2a - mous
7	26	39.4	39	2 T45362	acetyl-coenzyme a
8	26	39.4	41	2 G71376	hypothetical prote
9	26	39.4	47	2 T05922	hypothetical prote
10	26	39.4	50	2 AB0367	hypothetical prote
11	25	37.9	27	2 B42864	pancreatic cholest
12	25	37.9	30	2 S07217	ribosomal protein
13	25	37.9	32	2 S03352	malate dehydrogena
14	25	37.9	36	2 I57728	primase - Escheric
15	25	37.9	48	2 D90777	hypothetical prote
16	25	37.9	48	2 S42399	hypothetical prote
17	25	37.9	49	2 T07307	hypothetical prote
18	24	36.4	20	2 A56045	hypothetical prote
19	24	36.4	36	2 A60146	65K heat shock pro
20	24	36.4	39	2 F64843	hypothetical prote
21	24	36.4	41	2 S19600	hypothetical prote
22	24	36.4	43	2 S74097	AlsyI protein - mo
23	24	36.4	43	2 C84793	hypothetical prote
24	24	36.4	43	2 A84401	hypothetical prote
25	24	36.4	46	2 P00040	agrd protein - Sta
26	24	36.4	46	2 B45174	eye cell developme
27	24	36.4	47	2 PN0607	cytochrome-c oxida
28	23	34.8	20	2 T50757	pufk protein (impo
29	23	34.8	20	2 PN0459	basic eosinophil p

30	23	34.8	24	2 I67553	monocyte chemotact
31	23	34.8	27	2 A48200	homeotic protein H
32	23	34.8	35	2 B86327	protein F18014.13
33	23	34.8	38	2 S78757	ribosomal protein
34	23	34.8	41	2 T45906	hypothetical prote
35	23	34.8	41	2 S28769	hypothetical prote
36	23	34.8	43	2 D61219	serpin Spi 4 - hor
37	23	34.8	46	2 A70236	hypothetical prote
38	23	34.8	46	2 A60495	T-cell receptor be
39	23	34.8	47	2 B82245	hypothetical prote
40	23	34.8	47	2 A61048	hypothetical prote
41	23	34.8	48	2 I46522	tropoin T 2fa - r
42	22.5	34.1	23	4 A58505	interleukin-1 beta
43	22	33.3	14	2 C59137	protein P33 - gold
44	22	33.3	19	2 A37968	neural surface pro
45	22	33.3	24	2 S56003	glucan 1,3-beta-gl
46	22	33.3	25	2 S56002	glucan 1,3-beta-gl
47	22	33.3	28	2 A23691	apolipoprotein C-I
48	22	33.3	29	2 I84189	cyclic AMP recepto
49	22	33.3	31	2 C95215	hypothetical prote
50	22	33.3	31	2 C81193	hypothetical prote
51	22	33.3	31	2 E82446	hypothetical prote
52	22	33.3	32	2 A59156	gliadin omega-5 -
53	22	33.3	37	2 B85574	hypothetical prote
54	22	33.3	38	2 PH1920	annexin-like 40K p
55	22	33.3	40	2 T08870	alternative respir
56	22	33.3	40	2 A12387	photosystem I Psam
57	22	33.3	41	2 G82657	hypothetical prote
58	22	33.3	43	2 T07302	hypothetical prote
59	22	33.3	43	2 D97484	hypothetical prote
60	22	33.3	44	2 G82464	hypothetical prote
61	22	33.3	45	2 T52123	hypothetical prote
62	22	33.3	46	2 S14000	RNA-directed RNA p
63	22	33.3	46	2 S58637	hypothetical prote
64	22	33.3	46	2 B84115	hypothetical prote
65	22	33.3	47	2 AF3178	hypothetical prote
66	22	33.3	48	1 EMBY8	H+-transporting tw
67	22	33.3	48	1 PWC88P	H+-transporting tw
68	22	33.3	48	1 EMAS8	H+-transporting tw
69	22	33.3	48	1 EMAS8M	H+-transporting tw
70	22	33.3	48	2 S58742	H+-transporting tw
71	22	33.3	48	2 S78202	H+-transporting tw
72	22	33.3	48	2 S17994	H+-transporting tw
73	22	33.3	48	2 S15501	H+-transporting tw
74	22	33.3	50	2 G91003	host-nuclease inh1
75	22	33.3	50	2 A12434	hypothetical prote

## ALIGNMENTS

## RESULT 1

hypothetical protein BH1766 [imported] - Bacillus halodurans (strain C-125)  
F83870  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence-revision 01-Dec-2000 #text-change 15-Jun-2001  
R:Accession: F83870  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: F83870  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-29 <STO>  
A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA05485.1; GSPDB:G  
A:Experimental source: strain C-125  
A:Genetics:  
A:Gene: BH1766

Query Match 50.0%; Score 33; DB 2; Length 29;  
Best Local Similarity 62.5%; Pred. No. 21;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 FATELHER 12  
|||||:  
Db 19 FATQLHDK 26

RESULT 2  
S14101

apolipophorin III - migratory locust  
C:Species: Locusta migratoria (migratory locust)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Jul-1997  
C:Accession: S14101  
R:van der Horst, D.J.; van Doorn, J.M.; Voshol, H.; Kanost, M.R.; Ziegler, R.; Beenakker  
Eur. J. Biochem. 196, 509-517, 1991  
A:Title: Different isoforms of an apoprotein (apolipophorin III) associate with lipopro  
A:Reference number: S14101; MUID:9117026; PMID:2007409  
A:Accession: S14101  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-37 <HOR>

Query Match 43.9%; Score 29; DB 2; Length 37;  
Best Local Similarity 58.3%; Pred. No. 1.5e+02;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LVFPATELHERL 13  
:| | | | | |  
Db 20 IVNAHHEHETL 31

RESULT 3  
A61147

manganese peroxidase (EC 1.11.1.13), isoenzyme 4.9 - basidiomycete (Phanerochaete chrysoc  
C:Species: Phanerochaete chrysosporium  
C:Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 06-Dec-1996  
C:Accession: A61147  
R:Datla, A.; Blettermann, A.; Kirk, T.K.  
Appl. Environ. Microbiol. 57, 1453-1460, 1991  
A:Title: Identification of a specific manganese peroxidase among ligninolytic enzymes se  
A:Reference number: A61147; MUID:91307312; PMID:1854201  
A:Accession: A61147  
A:Molecule type: protein  
A:Residues: 1-40 <DAT>  
C:Superfamily: lignin peroxidase  
C:Keywords: extracellular protein; heme; manganese; oxidoreductase

Query Match 42.4%; Score 28; DB 2; Length 40;  
Best Local Similarity 54.5%; Pred. No. 2.5e+02;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 VPFAELHERL 13  
| | | | | |  
Db 18 VPLADLOETL 28

RESULT 4  
T52124

hypothetical protein MYB54 [imported] - Arabidopsis thaliana (fragment)  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000  
C:Accession: T52124  
R:Romero, I.; Fuentes, A.; Benito, M.J.; Malpica, J.; Leyva, A.; Paz-Ares, J.  
submitted to the EMBL Data Library, May 1997  
A:Description: One hundred R2R3-MYB genes in the genome of Arabidopsis thaliana.  
A:Reference number: Z25968  
A:Accession: T52124  
A:Status: preliminary;  
A:Molecule type: mRNA  
A:Residues: 1-45 <ROM>  
A:Cross-references: EMBL:Z95777; P1DN:CAB09209.1  
C:Genetics:  
A:Gene: MYB54

Query Match 42.4%; Score 28; DB 2; Length 45;  
Best Local Similarity 60.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
| | | | | |  
Db 15 PFTEEEERL 24

RESULT 5  
S15837

hypothetical protein (rep 3' region) - Rhodobacter sp. plasmid PRD31  
C:Species: Rhodobacter sp.  
C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 18-Jun-1993  
C:Accession: S15837  
R:Matunaga, T.; Mihashita, H.; Miyake, M.; Burgess, J.G.  
FEBS Lett. 283, 263-266, 1991  
A:Title: Nucleotide sequence of the replication region of the marine Rhodobacter plas  
A:Reference number: S15837; MUID:91257317; PMID:2044764  
A:Accession: S15837  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-31 <FEB>  
C:Genetics:  
A:Genome: plasmid

Query Match 39.4%; Score 26; DB 2; Length 31;  
Best Local Similarity 36.4%; Pred. No. 4.3e+02;  
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 LVFPATELHER 12  
:| | | | | |  
Db 10 WVPYLSQLEEQ 20

RESULT 6  
I48405

histone H2a - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999  
C:Accession: I48405  
R:Slitman, D.B.; Graves, R.A.; Marzluft, W.F.  
Nucleic Acids Res. 11, 6679-6697, 1983  
A:Title: Structure of a cluster of mouse histone genes.  
A:Reference number: I48401; MUID:84041477; PMID:6314253  
A:Accession: I48405  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-37 <RES>  
A:Cross-references: EMBL:X02622; MID:951322; P1DN:CAA26476.1; PID:951323  
C:Superfamily: histone H2A

Query Match 39.4%; Score 26; DB 2; Length 37;  
Best Local Similarity 45.5%; Pred. No. 5.2e+02;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 LVFPATELHER 12  
| | | | | |  
Db 23 LLPKTESHHK 33

RESULT 7  
T45362

acetyl-coenzyme A carboxylase carboxy transferase subunit beta [imported] - Mycobacte  
C:Species: Mycobacterium leprae  
C:Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 02-Sep-2000  
C:Accession: T45362  
R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, September 1997  
A:Reference number: Z16918  
A:Accession: T45362  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA

A:Residues: 1-39 <PAR>  
A:Cross-references: EMBL:Z99494; PIDN:CAB16708.1  
A:Experimental source: cosmid B57  
C:Genetics:  
A:Note: accD  
C:Superfamily: propionyl-CoA carboxylase beta chain

Query Match 39.4%; Score 26; DB 2; Length 39;  
Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 ATEHL 10  
|||  
DB 8 ATEHL 12

## RESULT 8

G1376  
Hypothetical protein TP0010 - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999

C:Accession: G1376

R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
rson, J.; Khailak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uitterback, T.; MCD  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A1250; MUID:98332770; PMID:9665876

A:Accession: G1376

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-41 <COLD>

A:Cross-references: GB:AE001186; GB:AE000520; NID:g3322263; PIDN:AC65011.1; PID:g332226  
A:Experimental source: strain Nichols

C:Genetics:  
A:Gene: TP0010

Query Match 39.4%; Score 26; DB 2; Length 41;  
Best Local Similarity 50.0%; Pred. No. 5.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 LVFPAELHE 11  
|||  
DB 20 VVFPVKRLSE 29

## RESULT 9

T05922  
Hypothetical protein - barley (fragment)

C:Species: Hordeum vulgare (barley)  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 08-Oct-1999

C:Accession: T05922

R:Hess, W.R.; Golz, R.R.; Boerner, T.  
Plant Sci. 133, 191-201, 1998

A:Title: Analysis of randomly selected cDNAs reveals the expression of stress- and defe  
A:Reference number: Z15411  
A:Accession: T05922

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA

A:Residues: 1-47 <HES>

A:Cross-references: EMBL:AJ222777; NID:e1203985; PIDN:CA10982.1; PID:e1203986  
A:Experimental source: cv. haisa, leaf

Query Match 39.4%; Score 26; DB 2; Length 47;  
Best Local Similarity 50.0%; Pred. No. 6.8e+02;  
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 LVFPAELHE 13  
|||  
DB 29 LVDFVMEIMSL 40

## RESULT 10

AB0367  
Hypothetical protein YPO3018 [Imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001

C:Accession: AB0367

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G  
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell  
Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AB0367

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-50 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC92261.1; PID:g15980972; GSPDB:GN00175

C:Genetics:  
A:Gene: YPO3018

Query Match 39.4%; Score 26; DB 2; Length 50;  
Best Local Similarity 41.7%; Pred. No. 7.2e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

OY 2 LVFPAELHE 11  
|||  
DB 14 LIPYSTRDLHE 25

## RESULT 11

B42864  
pancreatic cholesterol esterase homolog - human (fragment)

C:Species: Homo sapiens (man)  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999

C:Accession: B42864

R:Kumar, B.V.; Aleman-Gomez, J.A.; Colwell, N.; Lopez-Candales, A.; Bosner, M.S.; Spi  
Biochemistry 31, 6077-6081, 1992

A:Title: Structure of the human pancreatic cholesterol esterase gene.

A:Reference number: A42864; MUID:92329425; PMID:1627550

A:Accession: B42864

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-27 <KUM>

A:Cross-references: GB:S40179; NID:g251575; PIDN:AA822538.1; PID:g251576

A:Experimental source: SK-Hep-I, ATCC HTB 52

A:Note: sequence inconsistent with the nucleotide translation  
A:Note: sequence extracted from NCBI backbone (NCBIN:109186, NCBIIP:109187)

Query Match 37.9%; Score 25; DB 2; Length 27;  
Best Local Similarity 83.3%; Pred. No. 5.6e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 PFATPL 9  
|||  
DB 12 PFATPL 17

## RESULT 12

S07217  
ribosomal protein L18 [validated] - Halobacterium salinarum (fragment)

N:Alternate names: ribosomal protein HL13

C:Species: Halobacterium salinarum

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 21-Jul-2000

C:Accession: S07217; B33084; D33084

R:Smith, N.; Matheson, A.T.; Taguchi, M.; Willick, G.E.; Nazari, R.N.  
Eur. J. Biochem. 89, 501-509, 1978

A:Title: The 5-S RNA-protein complex from an extreme halophile, Halobacterium cutirub

A:Reference number: S07217; MUID:79045279; PMID:152199

A:Accession: S07217

A:Molecule type: protein

A:Residues: 1-30 <SMI>

A:Note: the source is designated as Halobacterium cutirubrum  
R:Mcdougall, J.

submitted to the Protein Sequence Database, June 1990

A:Reference number: A33084  
A:Accession: B33084  
A:Molecule type: protein  
A:Residues: 1-26 <MC2>  
A:Experimental source: strain DSM 670  
A>Note: The source is designated as Halobacterium halobium  
A:Accession: D33084  
A:Molecule type: protein  
A:Residues: 1-23 <MC2>  
A:Experimental source: strain DSM 3754  
A>Note: The source is designated as Halobacterium salinarum  
C:Complex: a 5S rRNA/protein complex contains ribosomal proteins L5 (PIR:S08569) and L18  
C:Superfamily: rat ribosomal protein L5  
C:Keywords: protein biosynthesis; ribosome

Query Match  
Best Local Similarity 37.9%; Score 25; DB 2; Length 30;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 TELHERL 13  
1 : 1 : 1 :  
DB 18 TDYHORL 24

## RESULT 13

S03352  
malate dehydrogenase (EC 1.1.1.37) - diatom (Nitzschia alba) (fragment)  
C:Species: Nitzschia alba  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 02-Aug-1994  
C:Accession: S03352  
R:Yueh, A.Y.; Chung, C.S.; Lai, Y.K.  
Biochem. J. 258, 221-228, 1989  
A:Title: Purification and molecular properties of malate dehydrogenase from the marine ciliate  
A:Reference number: S03352; MUID:89193527; PMID:2930509  
A:Accession: S03352  
A:Molecule type: protein  
A:Residues: 1-32 <YUE>  
C:Superfamily: L-lactate dehydrogenase  
C:Keywords: NAD; oxidoreductase

Query Match  
Best Local Similarity 37.9%; Score 25; DB 2; Length 32;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KLVPFATEL 9  
1 : 1 : 1 :  
DB 21 KLSQVTEEL 29

## RESULT 14

157728  
primase - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 08-Oct-1999  
C:Accession: 157728  
R:Lanka, E.; Lutz, R.; Kroger, M.; Furste, J.P.  
Mol. Gen. Genet. 194, 65-72, 1984  
A:Title: Plasmid RP4 encodes two forms of a DNA primase.  
A:Reference number: 157728; MUID:84218857; PMID:6374382  
A:Accession: 157728  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-36 <RES>  
A:Cross-references: EMBL:X00448; NID:g42506; PIDN:CAA25139.1; PID:g42507

Query Match  
Best Local Similarity 37.9%; Score 25; DB 2; Length 36;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 4 PFATELHERL 13  
1 : 1 : 1 :  
DB 7 PFHEQVAREL 16

## RESULT 15

D90777  
hypothetical protein ECs1188 [imported] - Escherichia coli (strain O157:H7, substrain  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C:Accession: D90777  
R:Hawash, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: D90777  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-48 <HAW>  
A:Cross-references: GB:BA000007; PIDN:BA834611.1; PID:g13360648; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain R1MD 0509952  
C:Genetics:  
A:Gene: ECs1188

Query Match  
Best Local Similarity 37.9%; Score 25; DB 2; Length 48;  
Matches 3; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 KLVPFATELHERL 13  
1 : 1 : 1 :  
DB 4 QLSFYQDKIRKHI 16

## RESULT 16

S42399  
hypothetical protein 48 - phage HK022  
C:Species: phage HK022  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 08-Oct-1999  
C:Accession: S42399  
R:Oberto, J.; Boehmer Sloan, S.; Weisberg, R.A.  
Nucleic Acids Res. 22, 354-356, 1994  
A:Title: A segment of the phage HK022 chromosome is a mosaic of other lambdaoid chromo  
A:Reference number: S42398; MUID:94173683; PMID:8127672  
A:Accession: S42399  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-48 <OBE>  
A:Cross-references: EMBL:U02466; NID:g407285; PIDN:AAB60271.1; PID:g407288

Query Match  
Best Local Similarity 37.9%; Score 25; DB 2; Length 48;  
Matches 3; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 KLVPFATELHERL 13  
1 : 1 : 1 :  
DB 4 QLSFYQDKIRKHI 16

## RESULT 17

T07307  
hypothetical protein 49d - Chlorella vulgaris chloroplast  
C:Species: chloroplast Chlorella vulgaris  
C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000  
C:Accession: T07307  
R:Makusagi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Na  
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga C  
A:Reference number: Z15985; MUID:97303241; PMID:9159184  
A:Accession: T07307  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-49 <MAK>  
A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BA57955.1; PID:g2224471  
C:Genetics:  
A:Genome: chloroplast



C:Keywords: chloroplast

Query Match 37.9%; Score 25; DB 2; Length 49;  
Best Local Similarity 66.7%; Pred. No. 1.1e+03;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LVPEPAT 7  
:||||:  
Db 40 IVPFST 45

RESULT 18

A56045

urinary tract stone matrix protein 12, 43k - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 12-Apr-1995 #sequence\_revision 12-Apr-1995 #text\_change 12-Apr-1995

C:Accession: A56045

R:Binette, J.P.; Binette, M.B.; Gawinowicz, M.A.; Kendrick, N.

submitted to the Protein Sequence Database, February 1995

A:Description: Isolation, characterization and sequence of stone proteins.

A:Reference number: A56046

A:Accession: A56045

A:Status: Preliminary

A:Molecule type: protein

A:Residues: 1-20 <BIN>

Query Match

Best Local Similarity 36.4%; Score 24; DB 2; Length 20;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LVPEPAT 8  
:||||:  
Db 1 MTPPWTE 7

RESULT 19

A60146

65k heat shock protein homolog - Bartonella bacilliformis (fragment)

C:Species: Bartonella bacilliformis

C>Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 12-Sep-1997

C:Accession: A60146

R:Knobloch, J.; Schreiber, M.

Am. J. Trop. Med. Hyg. 43, 373-379, 1990

A:Title: Bb65, a major immunoreactive protein of Bartonella bacilliformis.

A:Reference number: A60146; MUID:91052092; PMID:1700634

A:Accession: A60146

A:Molecule type: protein

A:Residues: 1-36 <KNO>

C:Comment: Bartonella bacilliformis is the causative agent of bartonellosis, an infectio

antigen for immune sera.

C:Superfamily: chaperonin groEL

C:Keywords: molecular chaperone

Query Match

Best Local Similarity 36.4%; Score 24; DB 2; Length 36;  
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 KIVPFATELHERL 13  
:||||:  
Db 3 KEVAFGNDARERM 15

RESULT 20

F64843

hypothetical protein b1016 - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002

C:Accession: F64843

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A:Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: F64843

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-39 <BLAT>

A:Cross-references: GB:AE000203; GB:U00096; NID:g1787248; PIDN:AACT4101.1; PID:g17872

A:Experimental source: strain K-12, substrain MG1655

Query Match

Best Local Similarity 36.4%; Score 24; DB 2; Length 39;  
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 VPFATELHERL 13  
:||||:  
Db 6 VPFILMREGI 16

RESULT 21

S19600

Alsy protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999

C:Accession: S19600

R:Kay, G.F.; Ashworth, A.; Penny, G.D.; Dunlop, M.; Swift, S.; Brockdorff, N.; Rastan

Nature 354, 486-489, 1991

A:Title: A candidate spermatogenesis gene on the mouse Y chromosome is homologous to

A:Reference number: S19600; MUID:92086049; PMID:1749428

A:Accession: S19600

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-41 <KAY>

A:Cross-references: GB:S70392; NID:9240453; PIDN:AAB20601.1; PID:g240454

C:Superfamily: ubiquitin-activating enzyme E1

Query Match

Best Local Similarity 36.4%; Score 24; DB 2; Length 41;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 FATELHERL 13  
:||||:  
Db 33 FGSDQERL 41

RESULT 22

S74097

protein-tyrosine kinase (EC 2.7.1.112) ITB, 38k - rat (fragments)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 11-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 04-Feb-2000

C:Accession: S74097

R:Brinati, A.M.; James, P.; Guerra, B.; Ruzzeno, M.; Donella-Deana, A.; Pinna, L.A.

Eur. J. Biochem. 240, 400-407, 1996

A:Title: The spleen protein-tyrosine kinase TPK-ITB is highly similar to the catalytic

A:Reference number: S74096; MUID:96439070; PMID:8841405

A:Accession: S74097

A:Molecule type: protein

A:Residues: 1-11;12-16;17-24;25-35;36-43 <BRU>

A:Experimental source: spleen

C:Superfamily: protein-tyrosine kinase ZAP-70; protein kinase homology; SH2 homology

C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

Query Match

Best Local Similarity 36.4%; Score 24; DB 2; Length 43;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 VPFATELHERL 11  
:||||:  
Db 3 LPMDTEVEYE 11

RESULT 23

C84793

hypothetical protein AC2937480 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: C84793  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon, L.;  
euser, D.; Nierman, M.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: C84793  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-43 <STO>  
A:Cross-references: GB:AE002093; NID:94056498; PIDN:AA098064.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g37480  
A:Map position: 2

Query Match 36.4%; Score 24; DB 2; Length 43;  
Best Local Similarity 41.7%; Pred. No. 1.4e+03;  
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 2 LVPATELHERL 13  
|:| | | |  
Db 14 LIPSLTPYHSG 25

RESULT 24  
A84401  
hypothetical protein Vng2511h [imported] - *Halobacterium* sp. NRC-1  
C:Species: *Halobacterium* sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: A84401  
R:Ng, W.V.; Kennedy, S.P.; Mahatras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
; Lettshauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hougl, D.W.; Maddocks, D.G.; Jaldic  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of *Halobacterium* species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: A84401  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-43 <STO>  
A:Cross-references: GB:AE004437; NID:910581904; PIDN:AA020573.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: VNG2511H

Query Match 36.4%; Score 24; DB 2; Length 43;  
Best Local Similarity 71.4%; Pred. No. 1.4e+03;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 TELHERL 13  
|:| | | |  
Db 25 TEIWERL 31

RESULT 25  
PQ0040  
agrd protein - *Staphylococcus aureus*  
C:Species: *Staphylococcus aureus*  
C:Date: 07-Sep-1990 #sequence\_revision 26-Apr-1996 #text\_change 15-Oct-1999  
C:Accession: S20796; PQ0040; S58479  
R:Kornblum, J.; Projan, S.J.; Kreiswirth, B.N.; Mogazeh, S.L.; Eisner, W.; Ross, H.; Nov  
submitted to the EMBL Data Library, March 1990  
A:Reference number: S20793  
A:Accession: S20796  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-46 <KOR>  
A:Cross-references: EMBL:X52543; NID:946505; PIDN:CAA36782.1; PID:946509  
R:Janzon, L.; Loeffel, S.; Arvidson, S.  
Mol. Gen. Genet. 219, 480-485, 1989  
A:Title: Identification and nucleotide sequence of the delta-lysin gene, hld, adjacent to  
A:Reference number: J00386; MUID:90158509; PMID:2622452

A:Accession: PQ0040  
A:Molecule type: DNA  
A:Residues: 1-45 <JAN>  
R:Novick, R.P.; Projan, S.J.; Kornblum, J.; Ross, H.F.; Ji, G.; Kreiswirth, B.; Vande  
Mol. Gen. Genet. 248, 446-458, 1995  
A:Title: The agr P2 operon: an autocatalytic sensory transduction system in *Staphylococcus aureus*.  
A:Reference number: S58478; MUID:96004766; PMID:7565609  
A:Accession: S58479  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-46 <NOV>  
A:Cross-references: EMBL:X52543; NID:946505; PIDN:CAA36782.1; PID:946509

Query Match 36.4%; Score 24; DB 2; Length 46;  
Best Local Similarity 80.0%; Pred. No. 1.5e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 7 TELHE 11  
|:| | | |  
Db 42 TOLHE 46

RESULT 26  
B45174  
eye cell development gene eya protein, splice form 2 - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C:Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 28-Feb-1997  
C:Accession: B45174  
R:Bonini, N.M.; Lelerson, W.M.; Benzer, S.  
Cell 72, 379-395, 1993  
A:Title: The eyes absent gene: genetic control of cell survival and differentiation 1  
A:Reference number: A45174; MUID:93161413; PMID:8431945  
A:Accession: B45174  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-46 <BON>  
A:Experimental source: adult head  
A>Note: sequence extracted from NCBI backbone (NCBIP:124851)  
C:Genetics:  
A:Gene: FLYbase:cl; eya; eyes absent  
A:Cross-references: FlyBase:FBgn0000320

Query Match 36.4%; Score 24; DB 2; Length 46;  
Best Local Similarity 66.7%; Pred. No. 1.5e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 TELHER 12  
|:| | | |  
Db 33 TDTHER 38

RESULT 27  
PN0607  
cytochrome-c oxidase (EC 1.9.3.1) chain VIIC - pig (fragment)  
C:Species: *Sus scrofa domestica* (domestic pig)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-May-1999  
C:Accession: PN0607  
R:Sillard, R.; Joenvall, H.; Mutt, V.  
Biochem. Biophys. Res. Commun. 195, 746-750, 1993  
A:Title: Characterization of porcine intestinal cytochrome c oxidase subunit VIIC, pu  
A:Reference number: PN0607; MUID:93384597; PMID:836926  
A:Accession: PN0607  
A:Molecule type: protein  
A:Residues: 1-47 <SIL>  
A:Experimental source: intestine  
C:Superfamily: cytochrome-c oxidase chain VIIC  
C:Keywords: mitochondrion; oxidoreductase; respiratory chain

Query Match 36.4%; Score 24; DB 2; Length 47;  
Best Local Similarity 46.2%; Pred. No. 1.6e+03;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 KLVFATELHERL 13



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 17:53:50 ; Search time 11 Seconds  
(Without alignments)  
49.017 Million cell updates/sec

Title: US-09-993-366-1  
Perfect score: 66  
Sequence: 1 KLVFPATELHERL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 3754

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	40.9	40	1	ALB1_TRASC
2	26	39.4	41	1	Y010_TREPA
3	25	37.9	26	1	RL18_HALHA
4	25	37.9	30	1	RL18_HALCU
5	25	37.9	32	1	MDH_NITVL
6	25	37.9	48	1	YO48_BPDKO
7	24	36.4	23	1	T2B_PARTE
8	23.5	35.6	28	1	COXB_SOLTU
9	23.5	35.6	33	1	FABI_RHASA
10	23	34.8	8	1	PPK3_PERAM
11	23	34.8	20	1	PUFK_RHOSH
12	23	34.8	23	1	CH60_THREE
13	23	34.8	24	1	DNAU_STRAG
14	23	34.8	35	1	PYC1_SPIPL
15	23	34.8	43	1	A1T4_HORSE
16	23	34.8	50	1	LHB2_RHOSH
17	23	34.8	50	1	LHB2_RHOSH
18	22	33.3	14	1	MARI_ALTSP
19	22	33.3	22	1	MOT1_CHICK
20	22	33.3	28	1	APC1_RABIT
21	22	33.3	31	1	GP37_BPSP1
22	22	33.3	32	1	CYSB_FASHE
23	22	33.3	42	1	LHB1_RHOAC
24	22	33.3	46	1	KRUP_LITFO
25	22	33.3	48	1	ATP8_ASFAM
26	22	33.3	48	1	ATP8_CANPA
27	22	33.3	48	1	ATP8_EMENI
28	22	33.3	48	1	ATP8_HANMI
29	22	33.3	48	1	ATP8_KIUIA
30	22	33.3	48	1	ATP8_SCHPO
31	22	33.3	48	1	ATP8_YARLI
32	22	33.3	48	1	ATP8_YEAST
33	22	33.3	50	1	ATP8_PODAN

34	22	33.3	50	1	Y0R1_TTV1	P19294 thermoprote
35	21.5	32.6	40	1	Y236_ARCFU	O30003 archaeoglob
36	21	31.8	22	1	IAD1_ENTFA	P24803 enterococc
37	21	31.8	23	1	RL18_HALVO	P50563 halobacteri
38	21	31.8	23	1	RL5_HALHA	P50566 halobacteri
39	21	31.8	23	1	T2A_PARTE	O27173 paramecium
40	21	31.8	24	1	PCW5_PACGO	P82457 pachycondyl
41	21	31.8	24	1	RS5_VIBPR	P52856 vibrio prot
42	21	31.8	25	1	COXO_ONCMY	P80334 oncohychnu
43	21	31.8	28	1	RL5_HALCU	P05972 halobacteri
44	21	31.8	28	1	UC05_MAIZE	P80611 zea mays (m
45	21	31.8	29	1	GALA_AMICA	P47214 amia calva
46	21	31.8	43	1	LHB2_RHOAC	P35095 rhodopsendo
47	21	31.8	46	1	PSAI_ANAVA	P58560 anabena sp
48	21	31.8	46	1	PSAI_ANAVA	P23079 anabena va
49	21	31.8	46	1	SRTA_STRPY	O9fdv1 streptococc
50	21	31.8	48	1	ATP8_CANPA	P05040 candida gla
51	21	31.8	48	1	ATP8_TRIRU	O36838 trichophyto
52	21	31.8	48	1	CYT_SOYBN	P25973 glycine max
53	21	31.8	48	1	YO48_BP22	O38661 bacterioph
54	21	31.8	8	1	PPK2_PERAM	P82692 periplaneta
55	20	30.3	12	1	XYLA_SPRVN	P14405 streptomyc
56	20	30.3	25	1	PCW2_PACGO	P82424 pachycondyl
57	20	30.3	31	1	CXOA_COMPE	P56712 conus penna
58	20	30.3	33	1	RL26_XENLA	P49629 xenopus lae
59	20	30.3	35	1	CP12_PIG	P80736 sus scrofa
60	20	30.3	36	1	PAHO_LARAR	P11337 latus argen
61	20	30.3	37	1	VG65_BPPH2	P16515 bacterioph
62	20	30.3	42	1	PETM_ODOST	P49531 odontella s
63	20	30.3	42	1	VP10_BPPH6	P11127 bacterioph
64	20	30.3	44	1	VG67_BPML5	O05281 mycobacteri
65	20	30.3	43	1	RPC3_BPDKO	P18681 bacterioph
66	20	30.3	45	1	FABP_YARLI	P80547 yarrowia li
67	20	30.3	46	1	KRUP_PHOPA	O02034 rhodospira pha
68	20	30.3	46	1	YN44_ARCFU	O30325 archaeoglob
69	20	30.3	48	1	LHB1_RHOCA	P02950 rhododacter
70	20	30.3	48	1	LHB1_RHOSH	P02951 rhododacter
71	20	30.3	49	1	Y234_HAEIN	P43968 haemophilus
72	20	30.3	49	1	LHB2_RHOCA	P07368 rhododacter
73	20	30.3	49	1	LHB_ROSDE	P26275 roseobacter
74	20	30.3	49	1	R22A_MOUSE	P35255 mus musculu
75	20	30.3	49	1	RL33_CIOAB	O97eg2 clostiridum

## ALIGNMENTS

RESULT 1  
ALB1\_TRASC STANDARD: PRT: 40 AA.  
AC P81188:  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE 67 kDa serum albumin (Alb-1) (Fragment).  
OS Trichemys scripta (Red-eared slider turtle) (Pseudemys scripta).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Testudines; Cryptodira; Testudinoidea; Emydidae; Trachemys.  
OX NCBI\_TaxID=34903;  
RN [1]  
RP MEDLINE=98103404; PubMed=9440230;  
RA Brown M.A., Chambers G.K., Licht P.;  
RT "Purification and partial amino acid sequences of two distinct  
RT albumins from turtle plasma.";  
RT Comp. Biochem. Physiol. 118B:367-374(1997).  
RL [2]  
RP MEDLINE=95309661; PubMed=7789749;  
RA Seitzer K.W., Palmer B.D.;  
RT "Estrogen downregulation of albumin and a 170-kDa serum protein in  
RT the turtle, Trichemys scripta.";  
RL Gen. Comp. Endocrinol. 97:340-352(1995).

CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD  
 CC BINDING CAPACITY FOR WATER, CA<sup>++</sup>, NA<sup>+</sup>, K<sup>+</sup>, FATTY ACIDS, HORMONES,  
 CC BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE  
 CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: PLASMA.  
 CC -1- MISCELLANEOUS: IN THE RED-EARED SLIDER TURTLE, THERE ARE TWO FORMS  
 CC OF ALBUMIN, ALB-1 AND ALB-2.  
 CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.  
 DR HSPSP: P02768; IE7H.  
 DR InterPro: IPR000264; Serum albumin.  
 DR Pfam: PF00223; transport\_prot.1.  
 DR PROSITE: PS00212; ALBUMIN, PARTIAL.  
 KW Plasma; Metal-binding; Lipid-binding; Albumin; Copper.  
 FT METAL 4  
 FT NON\_TER 40  
 FT SEQUENCE 40 AA; 4682 MW; 5FAC9E49E2789BB0 CRC64;

Query Match 40.98; Score 27; DB 1; Length 40;  
 Best Local Similarity 45.5%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LVFPATELHER 12  
 Db 25 LITFAOFLMK 35

RESULT 2  
 ID Y010.TREPA STANDARD: PRT; 41 AA.  
 AC 083054;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein TP0010.  
 GN TP0010.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxID=160;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NICHOLS;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Morris S.J., Weinstock G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
 RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 RT spirochete.";  
 RL Science 281:375-388(1998).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AE001186; AAC65011.1; -  
 DR TIGR: TP0010; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 41 AA; 4471 MW; 08178A6F95D83FB CRC64;

Query Match 39.48; Score 26; DB 1; Length 41;  
 Best Local Similarity 50.0%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Qy 2 LVFPATELHE 11  
 Db 11 LITFAOFLMK 11

Db 20 VVFPVKRLSE 29

RESULT 3  
 ID RL18\_HALHA STANDARD: PRT; 26 AA.  
 AC P50560;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 50S ribosomal protein L18P (HHA118) (Fragment).  
 GN RPL18P.  
 OS Halobacterium halobium.  
 OC Archaea; Euryarchaeota; Halobacteriales;  
 OC Halobacteriaceae; Halobacterium.  
 OX NCBI\_TaxID=2242;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=DSM 670;  
 RX MEDLINE=94229075; PubMed=8174557;  
 RA McDougall J., Wilmann-Liebold B.;  
 RT "Comparative analysis of the protein components from 5S rRNA-protein  
 RT complexes of halophilic archaeobacteria.";  
 RL Eur. J. Biochem. 221:779-785(1994).  
 CC -1- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.  
 DR PIR: B33084; B33084.  
 KW Ribosomal protein.  
 FT NON\_TER 26  
 FT SEQUENCE 26 AA; 3269 MW; D569C8746E1F607F CRC64;

Query Match 37.98; Score 25; DB 1; Length 26;  
 Best Local Similarity 57.18%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 7 TELHERL 13  
 Db 18 TDYHQR 24

RESULT 4  
 ID RL18\_HALCU STANDARD: PRT; 30 AA.  
 AC P05970;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 50S ribosomal protein L18P (HCU118) (HL13) (Fragment).  
 GN RPL18P.  
 OS Halobacterium cutribdium.  
 OC Archaea; Euryarchaeota; Halobacteriales;  
 OC Halobacteriaceae; Halobacterium.  
 OX NCBI\_TaxID=2242;  
 RN [1]  
 RP SEQUENCE.  
 RC MEDLINE=79045279; PubMed=152199;  
 RA Smith N., Matheson A.T., Taguchi M., Willick G., Nazar R.N.;  
 RT "The 5-S rRNA-protein complex from an extreme halophile,  
 RT Halobacterium cutribdium. Purification and characterization.";  
 RL Eur. J. Biochem. 89:501-509(1978).  
 CC -1- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.  
 DR PIR: S07217; S07217.  
 KW Ribosomal protein.  
 FT NON\_TER 30  
 FT SEQUENCE 30 AA; 3624 MW; 3A50079B1569CB74 CRC64;

Query Match 37.98; Score 25; DB 1; Length 30;  
 Best Local Similarity 57.18%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 7 TELHERL 13  
 Db 18 TDYHQR 24

```

RESULT 5
MDH_NITRAL
ID MDH_NITRAL STANDARD: PRT: 32 AA.
AC P10887;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Malate dehydrogenase (EC 1.1.1.37) (Fragment).
OS Nitzschia alba (Marine diatom).
OC Eukaryota; Stramenopiles; Bacillariophyta; Bacillariophyceae;
OC Bacillariophycidae; Bacillariales; Bacillariaceae; Nitzschia.
OX NCBI_TaxID=2858;
RN [1]
RP SEQUENCE.
RX MEDLINE=69193527; PubMed=2930509;
RA Yuen A.Y., Chung C.-S., Lai Y.-K.;
RT "Purification and molecular properties of malate dehydrogenase from
RL the marine diatom Nitzschia alba.";
RL Biochem. J. 258:221-228(1989).
CC -1- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.
CC -1- SUBUNIT: HOMODIMER (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
DR HSSP: S03352;
DR PIR: S03352;
DR HSSP: P00346; 1MD.
DR InterPro: IPR001252; Mdh.
DR InterPro: IPR001236; Ldh.
DR Pfam: PF00056; Ldh. 1.
DR PROSITE: PS00068; MDH; PARTIAL.
DR Oxidoreductase; Tricarboxylic acid cycle; NAD.
FT NON_TER 32
SQ SEQUENCE 32 AA: 3326 MW: 83F0E9E0BADB4C2 CRC64;

Query Match 37.9%; Score 25; DB 1; Length 32;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KLVFPATEL 9
Db 21 KLSQVTEL 29

RESULT 6
YO48_BPHKO
ID YO48_BPHKO STANDARD: PRT: 48 AA.
AC Q37928;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative 5.8 kDa protein (ORF48).
OS Bacteriophage HK022.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10742;
RN [1]
RP SEQUENCE FROM N.A.
RA Oberto J., Bohmer Sloan S., Weisberg R.A.;
RT "DNA sequence of the cII and replication genes of coliphage
RT HK022.";
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL: U02466; AAB60271.1; -
KW Hypothetical protein.
SQ SEQUENCE 48 AA: 5839 MW: 22DB5DE71A28C916 CRC64;

```

```

Query Match 37.9%; Score 25; DB 1; Length 48;
Best Local Similarity 23.1%; Pred. No. 4.6e+02;
Matches 3; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 KLVFPATELHERL 13
Db 4 QLSPYQKIKHKHI 16

RESULT 7
T2B_PARTE
ID T2B_PARTE STANDARD: PRT: 23 AA.
AC Q27174;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trichocyst matrix protein T2-B (Secretory granule protein T2-B)
DE (T2B 2-B) (Fragment).
GN T2B.
OS Paramacium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peritritida;
OC Paramacium.
OX NCBI_TaxID=5888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D4-2;
RX MEDLINE=96059477; PubMed=7579685;
RA Madeddu L., Gautier M.-C., Vayssie L., Houari A., Sperling L.;
RT "A large multigene family codes for the polypeptides of the
RT crystalline trichocyst matrix in Paramacium.";
RL Mol. Biol. Cell 6:649-659(1995).
RN [2]
RP PARTIAL SEQUENCE.
RC STRAIN=D4-2;
RX MEDLINE=95119139; PubMed=7819344;
RA Madeddu L., Gautier M.-C., le Caer J.P., de Loubresse N., Sperling L.;
RT "Protein processing and morphogenesis of secretory granules in
RT Paramacium.";
RL Biochimie 76:329-335(1994).
RL -1- FUNCTION: STRUCTURAL PROTEIN THAT CRYSTALLIZE INSIDE THE
CC TRICHOCYST MATRIX.
CC -1- SUBCELLULAR LOCATION: TRICHOCYST. THESE ARE ARCHITECTURALLY
CC COMPLEX SECRETORY STORAGE GRANULES-DOCKED AT THE PLASMA MEMBRANE,
CC READY TO RAPIDLY RESPOND TO AN EXOCYTIC STIMULUS.
CC -1- SIMILARITY: BELONGS TO THE T2B FAMILY.
CC -1- DATABASE: NAME-Protein Spotlight;
CC NOTE=Issue 3 of October 2000;
CC WWW="http://www.expasy.org/spotlight/articles/spllt003.html".
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL: U27510; AAA92610.1; -
KW Polypeptin; Structural protein; Multigene family.
FT NON_TER 23
FT NON_TER 23
SQ SEQUENCE 23 AA: 2706 MW: 184D4B1F2759BDB7 CRC64;

Query Match 36.4%; Score 24; DB 1; Length 23;
Best Local Similarity 41.7%; Pred. No. 3.3e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 KLVFPATELHER 12
Db 5 RLVSSTLDLDDR 16

RESULT 8

```

```

COXB_SOLITU STANDARD: PRT: 28 AA.
ID COXB_SOLITU
AC P80499;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cytochrome c oxidase polypeptide Vb (EC 1.9.3.1) (Fragment).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RC TISSUE: Tuber;
RC MEDLINE: 97077345; PubMed: 8919912;
RA Jansch L., Kruff V., Schmitz U.K., Braun H.P.;
RT "New insights into the composition, molecular mass and stoichiometry
RT of the protein complexes of plant mitochondria.";
RL Plant J. 9:357-368(1996).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) -> 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VB FAMILY.
DR InterPro: IPR002124; COX5B.
DR PROSITE: PS00848; COX5B; PARTIAL.
KW Oxidoreductases; Inner membrane; Mitochondrion.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3101 MW; 1EAF79E2682849C CRC64;

Query Match 35.6%; Score 23.5; DB 1; Length 28;
Best Local Similarity 54.5%; Pred. No. 5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 2 LVFATELHER 12
Db 13 VMPINTG-HER 22

RESULT 9
FABI_RHASA STANDARD: PRT: 33 AA.
ID FABI_RHASA
AC P81175;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE Fatty acid-binding protein, intestinal (I-FABP) (Fragments).
OS Rhamdia sapo.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Pinelodidae; Rhamdia.
OX NCBI_TaxID=55673;
RN [1]
RP SEQUENCE.
RC TISSUE: Intestine;
RC MEDLINE: 98036128; PubMed: 9370361;
RA di Pietro S.M., Dell'Angelica E.C., Veerkamp J.H., Sterin-Speziale N.,
RA Santome J.A.;
RT "Amino acid sequence, binding properties and evolutionary
RT relationships of the basic liver fatty-acid-binding protein from the
RT catfish Rhamdia sapo.";
RL Eur. J. Biochem. 249:510-517(1997).
CC -1- FUNCTION: FABP ARE THOUGHT TO PLAY A ROLE IN THE INTRACELLULAR
CC TRANSPORT OF LONG-CHAIN FATTY ACIDS AND THEIR ACYL-COA ESTERS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: INTESTINE.
CC -1- SIMILARITY: BELONGS TO THE FABP/P2/CRBP/CRABP FAMILY OF
CC TRANSPORTERS.
DR InterPro: IPR000463; Fatty_acid_BP.
DR PROSITE: PS00214; FABP; PARTIAL.
KW Transport; Lipid-binding.
FT NON_TER 1
FT NON_CONS 12 13

```

```

FT NON_CONS 20 21
FT NON_CONS 28 29
FT NON_TER 33 33
SQ SEQUENCE 33 AA; 3660 MW; 5BA16CC2880B7819 CRC64;

Query Match 35.6%; Score 23.5; DB 1; Length 33;
Best Local Similarity 41.7%; Pred. No. 5.9e+02;
Matches 5; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

QY 3 VPFA-TELHERL 13
Db 10 VPFSVSEIQRV 21

RESULT 10
PKK3_PERAM STANDARD: PRT: 8 AA.
ID PKK3_PERAM
AC P82618;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pyrokinin-3 (Pea-PK-3) (FXPRL-amide).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthoptera; Dictyoptera; Blattaria;
OC Blattodea; Blattellidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE: Retrocerebral complex;
RC MEDLINE: 99212469; PubMed: 10196736;
RA Predel R., Kellner R., Nachman R.J., Holman G.M., Rapus J., Gaege G.;
RT "Differential distribution of pyrokinin-isoforms in cerebral and
RT abdominal neurohemal organs of the American cockroach.";
RL Insect Biochem. Mol. Biol. 29:139-144(1999).
RN [2]
RP TISSUE SPECIFICITY.
RC MEDLINE: 20189894; PubMed: 10723010;
RA Predel R., Eckert M.;
RT "Tagma-specific distribution of FXPRLamides in the nervous system of
RT the American cockroach.";
RL J. Comp. Neurol. 419:352-363(2000).
CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC (MYOTROPIC ACTIVITY).
CC -1- TISSUE SPECIFICITY: CORPORA CARDIACA.
CC -1- MASS SPECTROMETRY: MW:996.5; METHOD: MALDI.
CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 8
FT SEQUENCE 8 AA; 997 MW; 0B34177409D772C7 CRC64;

Query Match 34.8%; Score 23; DB 1; Length 8;
Best Local Similarity 62.5%; Pred. No. 1.1e+05;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LVFATEL 9
Db 1 LVFERPRL 8

RESULT 11
PUFK_RHOSH STANDARD: PRT: 20 AA.
ID PUFK_RHOSH
AC Q53121; Q08033;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcriptional regulatory protein pufk.
OS PUFK.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;

```



```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RX MEDLINE-96349111: PubMed-8760918;
RA Cong L., Kaplan S.;
RT "Translational control of puf operon expression in Rhodobacter
  sphaeroides 2.4.1.";
RL Microbiology 142:2057-2069(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RA Mcglynn P.;
RT "R.sphaeroides genes bchC, bchX, bchY, bchZ and pufO.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RX MEDLINE-20115911: PubMed-10648776;
RA Choudhary M., Kaplan S.;
RT "DNA sequence analysis of the photosynthesis region of Rhodobacter
  sphaeroides 2.4.1.";
RL Nucleic Acids Res. 28:862-867(2000).
CC -!- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF PUF.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S82643: AAB6798.1: -
DR EMBL: AJ010302: CAB38751.1: -
DR EMBL: AF195122: AAF24301.1: -
RW Transcription regulation.
SQ SEQUENCE 20 AA; 2262 MW; 764DBD64B9D990C CRC64;

Query Match 34.8%; Score 23; DB 1; Length 20;
Best Local Similarity 25.0%; Pred. No. 4.4e+02;
Matches 3; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 LVFPATELHERL 13
Db 1 MVEPYRMRHQHV 12

RESULT 12
CH60_THIPE STANDARD; PRT; 23 AA.
AC P29134;
ID CH60_THIPE STANDARD; PRT; 23 AA.
DT 01-DEC-1992 (Rel. 24, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 60 kDa chaperonin (Protein Cpn60) (GroEL protein) (Fragment).
GN GROEL OR MOPA OR GROEL.
OS Thiobacillus ferrooxidans.
OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.
OX NCBI_TaxID-920;
RN [1]
RP SEQUENCE.
RC STRAIN-ATCC 19859;
RX MEDLINE-93093401: PubMed-1360930;
RA Varela P., Jerez C.A.;
RT "Identification and characterization of GroEL and DnaK homologues in
  Thiobacillus ferrooxidans.";
RL FEMS Microbiol. Lett. 77:149-153(1992).
RN [2]
RP PHOSPHORYLATION.
RX MEDLINE-96242312: PubMed-9026439;
RA Seeger M., Osorio G., Jerez C.A.;
RT "Phosphorylation of GroEL, DnaK and other proteins from Thiobacillus
  ferrooxidans grown under different conditions.";

```

```

RL FEMS Microbiol. Lett. 138:129-134(1996).
CC -!- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
CC CONDITIONS (BY SIMILARITY).
CC -!- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
CC 7 SUBUNITS.
CC -!- INDUCTION: BY HEAT SHOCK.
CC -!- PM: PHOSPHORYLATED ON THREONINE.
CC -!- MISCELLANEOUS: THIS PROTEIN SHOWS ATPASE ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
DR Prosite: P29134; -
DR InterPro: IPR001844; Chaperin_Cpn60.
DR PROSITE: PS00296; CHAPERONINS_CPN60; PARTIAL.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2536 MW; 2BB688387FEB255D CRC64;

Query Match 34.8%; Score 23; DB 1; Length 23;
Best Local Similarity 38.5%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 KLVFPATELHERL 13
Db 3 KOVAFAEHAREKM 15

RESULT 13
DNAJ_STRAG STANDARD; PRT; 24 AA.
ID DNAJ_STRAG STANDARD; PRT; 24 AA.
AC P95694;
ID P95694;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chaperone protein dnaJ (Fragment).
GN DNAJ.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID-1311;
RN [1]
RP SEQUENCE FROM N.A.
RA Rioux C.R., Martin D., Hamel J., Brodeur B.R.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
CC THE ATPASE ACTIVITY OF DNAK (BY SIMILARITY).
CC -!- COFACTOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE DNAJ FAMILY.
CC -!- SIMILARITY: CONTAINS 1 J DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U72719: AAB39220.1: -
DR InterPro: IPR001305; DnaJ_CXXCXXG.
DR InterPro: IPR001623; DnaJ_N.
DR Pfam: PF00226; DnaJ_1.
DR PROSITE: PS00636; DNAJ_1; PARTIAL.
DR PROSITE: PS50076; DNAJ_2; PARTIAL.
DR PROSITE: PS00637; DNAJ_CXXCXXG; PARTIAL.
KW Chaperone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding.
FT DOMAIN 3
FT NON_TER 24
SQ SEQUENCE 24 AA; 2760 MW; 342AE656E00913FC CRC64;

Query Match 34.8%; Score 23; DB 1; Length 24;
Best Local Similarity 57.1%; Pred. No. 5.3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

OY 7 TELHERL 13  
11:11  
Db 4 TEFYDRL 10

RESULT 14  
PCL\_SPIPL  
ID PCL\_SPIPL STANDARD: PRT: 35 AA.  
AC P72506;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DE 15-DEC-1998 (Rel. 37, Last annotation update)  
DE core (LC 7.8) (Fragment).  
GN APCC.  
OS Spirulina platensis.  
OC Bacteria: Cyanobacteria; Oscillatoriales; Spirulina.  
OX NCBI\_TaxID=1156;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN:Italy;  
RA Meesapyodasak D., Nomsasawal P., Anjard C., Tanticharoen M.,  
RU Cheevadhanarak S.;  
CC Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: THIS PROTEIN OCCURS IN THE CORE MEMBRANE  
CC OF PHYCOBILISOMES, ASSOCIATED WITH ALLOPHYCOCYANIN.  
CC -1- SIMILARITY: TO OTHER PHYCOBILISOME LINKER PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
CC -----  
DR EMBL: X95898; CA65143.1; -.  
DR HSSP: P20116; 1B33.  
DR InterPro: IPR001685; CPCD-1like.  
DR Pfam: PF01383; CPCD. 1.  
FT PHYCOBILISOME: Photosynthesis.  
FT NON\_TER 35  
SQ SEQUENCE 35 AA: 4205 MW: 8F8E5337CB3AB8AE CRC64;

Query Match 34.8%; Score 23; DB 1; Length 35;  
Best Local Similarity 80.0%; Pred. No. 7.8e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KLVPF 5  
11111  
Db 29 KLVVPY 33

RESULT 15  
ALT4\_HORSE  
ID ALT4\_HORSE STANDARD: PRT: 43 AA.  
AC P38031;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Alpha-1-antitrypsin 4 (Alpha-1-antitrypsin 4) (Alpha-1-proteinase  
DE inhibitor 4) (SPI4) (Fragments).  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE.  
RC TISSUE:Plasma;  
RX MEDLINE:92126011; PubMed:1772402;  
RA Patterson S.D., Bell K., Shaw D.C.;  
RT "The equine major plasma serpin multigene family: partial

RT characterization including sequence of the reactive-site regions.";  
RL Biochem. Genet. 29:477-499(1991).  
RN [2]  
RN STRUCTURE OF CARBOHYDRATES.  
RX MEDLINE:90267497; PubMed:2111994;  
RA Patterson S.D., Bell K.;  
RT "The carbohydrate side chains of the major plasma serpins of horse  
RT and wallaby: analyses of enzymatic and chemically treated (including  
RT Smith degradation) protein blots by lectin binding.";  
RL Biochem. Int. 20:429-436(1990).  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: PLASMA.  
CC -1- PTM: N-GLYCOSYLATED WITH CARBOHYDRATES HAVING BIANTENARY  
CC SIDE CHAINS.  
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.  
DR PIR: D61219; D61219.  
DR InterPro: IPR000215; Serpin.  
DR PROSITE: PS00284; SERPIN; 1.  
KW Serpin; Serine protease inhibitor; Glycoprotein; Plasma;  
KW Multigene family.  
FT NON\_CONS 19 20 REACTIVE BOND.  
FT ACT\_SITE 23 24  
FT NON\_TER 43 43  
SQ SEQUENCE 43 AA: 4925 MW: 18E9430C1F4A87A CRC64;

Query Match 34.8%; Score 23; DB 1; Length 43;  
Best Local Similarity 33.3%; Pred. No. 9.6e+02;  
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 PFATELHER 12  
11:11  
Db 35 PFVLSTYDR 43

RESULT 16  
LHB2\_RHOSH  
ID LHB2\_RHOSH STANDARD: PRT: 50 AA.  
AC P02952;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Light-harvesting protein B-800/850, beta chain (LH-3B) (Antenna  
DE pigment protein, beta chain).  
GN PUCB.  
OS Rhodospirillum rubrum (Rhodospirillum rubrum).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum  
OC Rhodospirillum.  
OX NCBI\_TaxID=1063;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE:87250299; PubMed:3036782;  
RA Kiley P.J., Kaplan S.;  
RT "Cloning, DNA sequence, and expression of the Rhodospirillum  
RT light-harvesting B800-850-alpha and B800-850-beta genes.";  
RL J. Bacteriol. 169:3268-3275(1987).  
RN [2]  
RN SEQUENCE FROM N.A.  
RA Ashby M.K., Coomber S.A., Hunter C.N.;  
RT "Cloning, nucleotide sequence and transfer of genes for the B800-850  
RT light harvesting complex of Rhodospirillum rubrum.";  
RL FEBS Lett. 213:245-248(1987).  
RN [3]  
RN SEQUENCE FROM N.A.  
RX MEDLINE:90132586; PubMed:2693605;  
RA Burgess J.G., Ashby M.K., Hunter C.N.;  
RT "Characterization and complementation of a mutant of Rhodospirillum  
RT rubrum with a chromosomal deletion in the light-harvesting (LH2)  
RT genes.";  
RL J. Gen. Microbiol. 135:1809-1816(1989).  
RN [4]  
RN SEQUENCE FROM N.A.  
RX STRAIN:ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;  
RA MEDLINE:93086425; PubMed:1453956;

RA Gibson L.C., McGlynn P., Chaudhuri M., Hunter C.N.;  
 RT "A putative anaerobic coproporphyrinogen III oxidase in Rhodobacter  
 RT sphaeroides. II. Analysis of a region of the genome encoding heme f and  
 RT the puc operon."; Mol. Microbiol. 6:3171-3186(1992).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;  
 RX MEDLINE=20115911; PubMed=1064876;  
 RA Choudhary M., Kaplan S.;  
 RT "DNA sequence analysis of the photosynthesis region of Rhodobacter  
 RT sphaeroides 2.4.1."; Nucleic Acids Res. 28:862-867(2000).  
 RL [6]  
 RN SEQUENCE (MUTANT R-26.1).  
 RX MEDLINE=85005243; PubMed=6384009;  
 RA Thellier R., Suter F., Wiemken V., Zuber H.;  
 RT "The light-harvesting polypeptides of Rhodospseudomonas sphaeroides R-  
 RT 26.1. I. Isolation, purification and sequence analyses."; Hoppe-Seyler's Z. Physiol. Chem. 365:703-719(1984).  
 RL [7]  
 CC -1- FUNCTION: ANTENNA COMPLEXES ARE LIGHT-HARVESTING SYSTEMS, WHICH  
 CC TRANSFER THE EXCITATION ENERGY TO THE REACTION CENTERS.  
 CC -1- SUBUNIT: THE CORE COMPLEX IS FORMED BY DIFFERENT ALPHA AND BETA  
 CC CHAINS, BINDING BACTERIOCHLOROPHYLL MOLECULES, AND ARRANGED  
 CC MOST PROBABLY IN TETRAMERIC STRUCTURES DISPOSED AROUND THE  
 CC REACTION CENTER. THE NON-PIGMENTED GAMMA CHAINS MAY CONSTITUTE  
 CC ADDITIONAL COMPONENTS.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.  
 CC -1- MISCELLANEOUS: THIS POLYPEPTIDE (LH-3b) AND LH-2 CONSTITUTE THE B-  
 CC 800/850 COMPLEX OF R.SPHEROIDES 2.4.1 AND THE SPECTRALLY ALTERED  
 CC B-850 COMPLEX ISOLATED FROM THE BLUE-GREEN MUTANT R-26.1, WHICH  
 CC ABSORBS AT 860 NM.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M16777; AAA26159.1; -;  
 DR EMBL: X05200; CAA28832.1; -;  
 DR EMBL: M28360; AAA26130.1; -;  
 DR EMBL: X68796; CAA48699.1; -;  
 DR EMBL: AF195122; AAF24247.1; -;  
 DR PIR: A03454; LBRP85;  
 DR PIR: B27087; B27087;  
 DR PIR: A26645; A26645;  
 DR PIR: A45795; A45795;  
 DR PIR: S28023; S28023;  
 DR HSSP: P26790; 1KZU.  
 DR InterPro: IPR000066; Antenna\_a/b.  
 DR InterPro: IPR002362; Antenna\_beta.  
 DR Pfam: PF00556; LHC; 1.  
 DR PRINTS: PR00674; LIGHTHARVSTB.  
 DR PROSITE: PS00969; ANTENNA\_COMP\_BETA; 1.  
 DR Antenna complex; Light-harvesting polypeptide; Transmembrane;  
 KW Magnesium; Bacteriochlorophyll; Inner membrane.  
 FT INIT\_MET 0  
 FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 23 45 POTENTIAL.  
 FT DOMAIN 46 50 PERIPLASMIC (POTENTIAL).  
 FT BINDING 21 21 BACTERIOCHLOROPHYLL (POTENTIAL).  
 FT METAL 39 39 AXIAL LIGAND TO THE BACTERIOCHLOROPHYLL  
 FT CONFLICT 37 37 MAGNESIUM (POTENTIAL).  
 FT I -> L (IN REF. 2).  
 SO SEQUENCE 50 AA: 5317 MW: 103613DA59CEAF2 CRC64;  
 Query Match 34.8%; Score 23; DB 1; Length 50;  
 Best Local Similarity 50.0%; Pred. No. 1.le+03;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 6 ATTELHERL 13  
 11:11:1  
 Db 17 AEEVHKOL 24  
 RESULT 17  
 LHB2\_RHOSU  
 ID LHB2\_RHOSU STANDARD: PRT: 50 AA.  
 AC P95654;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Light-harvesting protein B-800/850, beta chain (LH-3b) (Antenna  
 DE pigment protein, beta chain).  
 DE PUBB  
 OS Rhodovulum sulfidophilum (Rhodobacter sulfidophilus).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;  
 OC Rhodovulum.  
 OX NCBI\_TaxID=35806;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=W4;  
 RX MEDLINE=97276905; PubMed=9130598;  
 RA Hagemann G.E., Katsiou E., Forkl H., Steindorf A.C., Tadros M.H.;  
 RT "Gene cloning and regulation of gene expression of the puc operon  
 RT from Rhodovulum sulfidophilum."; Biochim. Biophys. Acta 1351:341-358(1997).  
 RL [2]  
 RN SEQUENCE.  
 RP STRAIN=W4;  
 RC MEDLINE=95354845; PubMed=7628614;  
 RX Tadros M.H., Hagemann G.E., Katsiou E., Dierstein R., Schiltz E.;  
 RA "Isolation and complete amino acid sequence of the beta- and alpha-  
 RT polypeptides from the peripheral light-harvesting pigment-protein  
 RT complex II of Rhodobacter sulfidophilus."; FEBS Lett. 368:243-247(1995).  
 RL [3]  
 CC -1- FUNCTION: ANTENNA COMPLEXES ARE LIGHT-HARVESTING SYSTEMS, WHICH  
 CC TRANSFER THE EXCITATION ENERGY TO THE REACTION CENTERS.  
 CC -1- SUBUNIT: THE CORE COMPLEX IS FORMED BY DIFFERENT ALPHA AND BETA  
 CC CHAINS, BINDING BACTERIOCHLOROPHYLL MOLECULES, AND ARRANGED  
 CC MOST PROBABLY IN TETRAMERIC STRUCTURES DISPOSED AROUND THE  
 CC REACTION CENTER. THE NON-PIGMENTED GAMMA CHAINS MAY CONSTITUTE  
 CC ADDITIONAL COMPONENTS.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U81968; AAB59006.1; -;  
 DR HSSP: P26790; 1KZU.  
 DR InterPro: IPR000066; Antenna\_a/b.  
 DR InterPro: IPR002362; Antenna\_beta.  
 DR Pfam: PF00556; LHC; 1.  
 DR PRINTS: PR00674; LIGHTHARVSTB.  
 DR PROSITE: PS00969; ANTENNA\_COMP\_BETA; 1.  
 DR Antenna complex; Light-harvesting polypeptide; Transmembrane;  
 KW Magnesium; Bacteriochlorophyll; Inner membrane.  
 FT INIT\_MET 0  
 FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 23 45 POTENTIAL.  
 FT DOMAIN 46 50 PERIPLASMIC (POTENTIAL).  
 FT BINDING 21 21 BACTERIOCHLOROPHYLL (POTENTIAL).  
 FT METAL 39 39 AXIAL LIGAND TO THE BACTERIOCHLOROPHYLL  
 FT CONFLICT 37 37 MAGNESIUM (POTENTIAL).  
 FT A96F2B914D757661 CRC64;  
 SO SEQUENCE 50 AA: 5438 MW: A96F2B914D757661 CRC64;  
 Query Match 34.8%; Score 23; DB 1; Length 50;  
 Best Local Similarity 50.0%; Pred. No. 1.le+03;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 ATELHERL 13  
| : : : |  
Db 17 AEEVHKOL 24

RESULT 18  
MARI\_ALTPSP  
ID MARI\_ALTPSP STANDARD; PRT; 14 AA.  
AC P29359;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Marinstatin C-2 (Marinostatin C-1; Marinostatin D).  
OS Alteromonas sp. (strain B-10-31).  
OC Bacteria: Proteobacteria; gamma subdivision; Alteromonadaceae;  
OC Alteromonas.  
OX NCBI\_TaxID:29456;  
RN [1]  
RP SEQUENCE, AND ACTIVE SITE.  
RX MEDLINE:92176155; PubMed:1794974;  
RA Takano R., Imada C., Kamai K., Hara S.;  
RT "The reactive site of marinostatin, a proteinase inhibitor from  
RL marine Alteromonas sp. B-10-31.";  
J. Biochem. 110:856-858(1991).  
CC -1- FUNCTION: INHIBITS SUBTILISIN, CHYMOTRYPSIN, AND ELASTASE, BUT  
NOT TRYPSIN.  
KW Serine protease inhibitor.  
FT PEPTIDE 1 14 MARINOSTATIN C-2.  
FT PEPTIDE 3 14 MARINOSTATIN C-1.  
FT PEPTIDE 4 14 MARINOSTATIN D.  
FT MOD.RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT ACT\_SITE 6 7 REACTIVE BOND.  
SQ SEQUENCE 14 AA; 1644 MW; 6E7CEE92EF32E44 CRC64;

Query Match 33.3%; Score 22; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PFAT 7  
| | | |  
Db 2 PFAT 5

RESULT 19  
MOTL\_CHICK  
ID MOTL\_CHICK STANDARD; PRT; 22 AA.  
AC O9PBP6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Motilin.  
GN MLN.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID:9031;  
RN [1]  
RP SEQUENCE.  
RC TISSUE: Small intestine;  
RX MEDLINE:96255357; PubMed:8801522;  
RA De Clercq P., Depoortere I., Maclellan M., Vandermeers A.,  
RA Vandermeers-Piret M.C., Peeters T.L.;  
RT "Isolation, sequence, and bioactivity of chicken motilin.";  
RL Peptides 17:203-208(1996).  
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE REGULATION OF  
CC INTERDIGESTIVE GASTROINTESTINAL MOTILITY AND INDIRECTLY CAUSES  
CC RHYTHMIC CONTRACTION OF DUODENAL AND COLONIC SMOOTH MUSCLE.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: BELONGS TO THE MOTILIN FAMILY.  
KW Hormone.

SQ SEQUENCE 22 AA; 2686 MW; 6FD6A79B7FB02680 CRC64;

Query Match 33.3%; Score 22; DB 1; Length 22;  
Best Local Similarity 66.7%; Pred. No. 7.4e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VPFAE 8  
| | | |  
Db 2 VPFAE 7

RESULT 20  
APOC1\_RABIT  
ID APOC1\_RABIT STANDARD; PRT; 28 AA.  
AC P33047;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Apolipoprotein C-I (Apo-CI) (Fragment).  
GN APOC1.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID:9986;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE:91093092; PubMed:2266137;  
RA Weisgraber K.H., Mahley R.W., Kowal R.C., Herz J., Goldstein J.L.,  
RA Brown M.S.;  
RT "Apolipoprotein C-I modulates the interaction of apolipoprotein E  
RT with beta-migrating very low density lipoproteins (beta-VLDL) and  
RT inhibits binding of beta-VLDL to low density lipoprotein  
RT receptor-related protein.";  
RL J. Biol. Chem. 265:22453-22459(1990).  
CC -1- FUNCTION: APPEARS TO MODULATE THE INTERACTION OF APOE WITH BETA-  
CC MIGRATING VLDL AND INHIBIT BINDING OF BETA-VLDL TO THE LDL  
CC RECEPTOR-RELATED PROTEIN.  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: SECRETED IN PLASMA.  
CC -1- SIMILARITY: BELONGS TO THE APOC1 FAMILY.  
DR PIR: A23691; A23691.  
DR HSSP: P02654; IALE.  
KW Plasma; Lipid transport; VLDL.  
FT NON\_TER 28  
SQ SEQUENCE 28 AA; 3165 MW; 2903C027F676541C CRC64;

Query Match 33.3%; Score 22; DB 1; Length 28;  
Best Local Similarity 41.7%; Pred. No. 9.5e+02;  
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 KVPFAEHLER 12  
| | | | |  
Db 13 KLEFGNTLEEK 24

RESULT 21  
GP37\_BPSP1  
ID GP37\_BPSP1 STANDARD; PRT; 31 AA.  
AC 048393;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Putative gene 37 protein.  
GN 37.  
OS Bacteriophage SP01.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
OC SP01-like viruses.  
OX NCBI\_TaxID:10685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE:96327781; PubMed:9657951;  
RA Stewart C.R., Gashlightwala I., Hinata K., Krollkowski K.A.,  
RA Needleman D.S., Peng A.S.-Y., Peterman M.A., Tobias A., Wei P.;

```

RT "Genes and regulatory sites of the 'host-takeover module' in the
RT terminal redundancy of Bacillus subtilis bacteriophage SPOL."
RL Virology 246:329-340(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF031901; AAC29006.1; .
KW Hypothetical protein.
SQ SEQUENCE 31 AA; 3715 MW; 5ECBA628AF320670 CRC64;

Query Match 33.3%; Score 22; DB 1; Length 31;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LVFPAT 7
Db 5 IIPFLT 10

RESULT 22
CYSB_FASHE STANDARD; PRT; 32 AA.
ID CYSB_FASHE
AC P80529; P80531;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cathepsin B-like cysteine proteinase (EC 3.4.22.-) (Newly excysted
DE juvenile proteins 5 and 7) (Fragments).
OS Fasciola hepatica (liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
OC Echinostomida; Echinostomata; Fasciolidae; Fasciola.
OX NCBI_TaxID=6192;
RN [1]
RP MEDLINE=95366993; PubMed=7639732;
RA Kraljevic J., Ashman K., Meusen E.;
RT "Fasciola hepatica: rapid identification of newly excysted juvenile
RT proteins."
RL Biochem. Biophys. Res. Commun. 213:169-174(1995).
CC -1- FUNCTION: THIOL PROTEASE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT THE NEWLY EXCYSTED JUVENILE
CC STAGE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
DR InterPro: IPR000169; SHProl_acsile.
DR PROSITE: PS00139; THIOL_PROTEASE_CYS; PARTIAL.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS; PARTIAL.
DR PROSITE: PS00640; THIOL_PROTEASE_ASN; PARTIAL.
KW Hydrolyase; Thiol protease; zymogen.
FT PROPEP 1 >22 ACTIVATION PEPTIDE.
FT NON_CONS 22 >23
FT CHAIN <23 >32 CATHEPSIN B-LIKE CYSTEINE PROTEINASE.
FT NON_TER 32 >32
SQ SEQUENCE 32 AA; 3940 MW; 94BCD2E57BE95CFE CRC64;

Query Match 33.3%; Score 22; DB 1; Length 32;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 PFATL 9
Db 10 PFSDEL 15

RESULT 23
LHBI_RHOAC STANDARD; PRT; 42 AA.
ID LHBI_RHOAC
AC P35094;

```

```

DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Light-harvesting protein B-800/820, beta chain (antenna pigment
DE protein, beta chain).
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Rhodoblastus.
OX NCBI_TaxID=1074;
RN [1]
RP SEQUENCE.
RC STRAIN=7050;
RA Brunschwiler R.A., Bissig I., Niederer E., Suter F., Zuber H.;
RL (In) Biggins J. (eds.);
RL Progress in photosynthesis research, pp.11.1:13-16, Martinus Nijhoff,
RL The Hague (1987).
CC -1- FUNCTION: ANTENNA COMPLEXES ARE LIGHT-HARVESTING SYSTEMS, WHICH
CC TRANSFER THE EXCITATION ENERGY TO THE REACTION CENTERS.
CC -1- SUBUNIT: THE CORE COMPLEX IS FORMED BY DIFFERENT ALPHA AND BETA
CC CHAINS, BINDING BACTERIOCHLOROPHYLL MOLECULES, AND ARRANGED
CC MOST PROBABLY IN TETRAMERIC STRUCTURES DISPOSED AROUND THE
CC REACTION CENTER. THE NON-PIGMENTED GAMMA CHAINS MAY CONSTITUTE
CC ADDITIONAL COMPONENTS.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
CC HSSP: P26790; 1KZU.
DR InterPro: IPR000066; Antenna_a/b.
DR InterPro: IPR002362; Antenna_beta.
DR Pfam: PF00556; LHC; 1.
DR PRINTS: PR00674; LIGHTHARVSTB.
DR PROSITE: PS00969; ANTENNA_COMP_BETA; 1.
KW Antenna complex; Light-harvesting polypeptide; Transmembrane;
KW Magnesium; Bacteriochlorophyll; Inner membrane.
FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
FT TRANSFM 15 37 POTENTIAL.
FT DOMAIN 38 42 PERIPLASMIC (POTENTIAL).
FT BINDING 13 13 BACTERIOCHLOROPHYLL (POTENTIAL).
FT METAL 31 31 AXIAL LIGAND TO THE BACTERIOCHLOROPHYLL
FT FT MAGNESIUM (POTENTIAL).
SQ SEQUENCE 42 AA; 4725 MW; 11512F5CD7A008DE CRC64;

Query Match 33.3%; Score 22; DB 1; Length 42;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 6 ATELHERL 13
Db 9 ADELKHY 16

RESULT 24
KRUP_LITFO STANDARD; PRT; 46 AA.
ID KRUP_LITFO
AC Q01872;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Krueppel protein (Fragment).
GN KR.
OS Litobius forficatus.
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Myriapoda; Chilopoda;
OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
OX NCBI_TaxID=7552;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93066327; PubMed=1438276;
RX Sommer R.J., Retzlaff M., Goerlich K., Sander K., Tautz D.;
RT "Evolutionary conservation pattern of zinc-finger domains of
RT Drosophila segmentation genes."
RL Proc. Natl. Acad. Sci. U.S.A. 89:10782-10786(1992).
CC -1- FUNCTION: KRUEPPEL IS A GAP CLASS SEGMENTATION PROTEIN.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.

```

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L01598: AAA29280.1; -;  
 DR InterPro: IPR000823: Znf\_C2H2.  
 DR Pfam: PF00096: zfc2h2.1.  
 DR SMART: SM00355: Znf\_C2H2; 1.  
 DR PROSITE: PS00028: ZINC\_FINGER\_C2H2\_1; 1.  
 DR PROSITE: PS50157: ZINC\_FINGER\_C2H2\_2; 1.  
 KW Developmental protein; Gap protein; Zinc-finger;  
 KW Metal-binding; DNA-binding; Repeat; Nuclear protein.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 >46 ZINC-FINGERS.  
 FT ZN\_FING <1 4 C2H2-TYPE.  
 FT ZN\_FING 10 32 C2H2-TYPE.  
 FT ZN\_FING 38 >46 C2H2-TYPE.  
 FT NON\_TER 46 46  
 SQ SEQUENCE 46 AA: 5493 MW: FEE4AF3297A72522 CRC64:

Query Match 33.3%; Score 22; DB 1; Length 46;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 PRATEL 9  
 1:111  
 DB 37 PYACEL 42

RESULT 25  
 ID ATP8\_ASPAM STANDARD: PRT: 48 AA.  
 AC P00858;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).  
 GN ATP8.  
 OS Aspergillus amstelodami.  
 OG Mitochondrion.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 CC Eurotiales; Trichocomaceae; Eurotium.  
 OX NCBI\_TaxID=5054;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lazarus C.M., Kuntzel H.;  
 RL Submitted (MAR-1984) to the PIR data bank.  
 CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT  
 CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2O) + H(+)(in) = ADP + phosphate +  
 CC H(+)(out).  
 CC -1- SUBCELLULAR LOCATION: Membrane-bound.  
 CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.  
 CC PIR: A01069; EMBL: EMBL.  
 KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.  
 GN SEQUENCE 48 AA: 5772 MW: 2E514749C973D7CC CRC64;

Query Match 33.3%; Score 22; DB 1; Length 48;  
 Best Local Similarity 80.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KLVPF 5  
 :1111  
 DB 3 OLVPF 7

RESULT 26  
 ATP8\_CANPA

ID ATP8\_CANPA STANDARD: PRT: 48 AA.  
 AC p17345;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).  
 GN ATP8 OR AAP1.  
 OS Candida parapsilosis (Yeast).  
 OG Mitochondrion.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; microsporid Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5480;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS 7154 / SP1;  
 RA MEDLINE=90332440; PubMed=2143015;  
 RX Guelin E., Velours J., Guerin M.;  
 RT "Cloning and sequencing of a fragment of the linear mitochondrial DNA  
 RT of FO ATP synthase and a putative t-RNA(Pro).";  
 RL Nucleic Acids Res. 18:4267-4267(1990).  
 CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT  
 CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2O) + H(+)(in) = ADP + phosphate +  
 CC H(+)(out).  
 CC -1- SUBCELLULAR LOCATION: Membrane-bound.  
 CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X52115: CAA36361.1; -;  
 DR PIR: S10465; PWCX8P.  
 KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.  
 FT TRANSMEM 12 POTENTIAL.  
 SQ SEQUENCE 48 AA: 5492 MW: DE7876341CAF0747 CRC64;

Query Match 33.3%; Score 22; DB 1; Length 48;  
 Best Local Similarity 80.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KLVPF 5  
 :1111  
 DB 3 OLVPF 7

RESULT 27  
 ID ATP8\_EMENI STANDARD: PRT: 48 AA.  
 AC P00857;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).  
 GN ATP8.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OG Mitochondrion.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 CC Eurotiales; Trichocomaceae; Emericella.  
 OX NCBI\_TaxID=5072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83038633; PubMed=6290989;  
 RA Netzer R., Koehler H.G., Basak N., Kuntzel H.;  
 RT "Nucleotide sequence of Aspergillus nidulans mitochondrial genes  
 RT coding for ATPase subunit 6, cytochrome oxidase subunit 3, seven  
 RL unidentified proteins, four tRNAs and L-rRNA.";  
 Nucleic Acids Res. 10:4783-4794(1982).

[2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE-82247225; PubMed-6285306;  
RA Grisi E., Brown T.A., Waring R.B., Scanzocchio C., Davies R.W.;  
RT "Nucleotide sequence of a region of the mitochondrial genome of  
RL Aspergillus nidulans including the gene for ATPase subunit 6.>";  
Nucleic Acids Res. 10:3531-3539(1982).  
CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT  
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.  
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +  
CC H(+) (out).  
CC -1- SUBCELLULAR LOCATION: Membrane-bound.  
CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: J01390; AAA99204.1; -;  
DR EMBL: X04161; CAA27772.1; -;  
DR EMBL: X01507; CAA25708.1; -;  
DR PIR: A01068; EMAS8.  
KM Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.  
SQ SEQUENCE 48 AA; 5770 MW; A5CA50379097D5EF CRC64;  
  
Query Match 33.3%; Score 22; DB 1; Length 48;  
Best Local Similarity 80.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 KLVPE 5  
Db 3 QLVPE 7  
  
RESULT 28  
ATP8\_HANWI STANDARD; PRT; 48 AA.  
AC P48882;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).  
GN ATP8.  
OS Hansenula winekei (Yeast).  
OG Mitochondrion.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Pichia.  
OX NCBI\_Taxid=4907;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=21;  
RA Sekito T., Okamoto K., Kilano H., Yoshida K.;  
RT "Hansenula winekei mitochondrial genome's complete DNA sequence  
RT demonstrated unique characteristics.";  
RL Nucleic Acids Symp. Ser. 31:233-234(1994).  
CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT  
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.  
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +  
CC H(+) (out).  
CC -1- SUBCELLULAR LOCATION: Membrane-bound.  
CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC -----  
DR EMBL: D31785; BAA06565.1; -;  
KM Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.  
FT TRANSMEM POTENTIAL.  
SQ SEQUENCE 48 AA; 5879 MW; 2714F6968D88BC CRC64;  
  
Query Match 33.3%; Score 22; DB 1; Length 48;  
Best Local Similarity 80.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 KLVPE 5  
Db 3 QLVPE 7  
  
RESULT 29  
ATP8\_KLUOA STANDARD; PRT; 48 AA.  
ID ATP8\_KLUOA  
AC Q00608;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).  
GN ATP8.  
OS Kluyveromyces fragilis (Yeast).  
OG Mitochondrion.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
OX NCBI\_Taxid=28985;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K8;  
RA Hardy C.M., Clark-Walker G.D.;  
RT "Nucleotide sequence of the COX1 gene in Kluyveromyces fragilis  
RT mitochondrial DNA: evidence for recent horizontal transfer of a group  
RT II intron.";  
RL Curr. Genet. 20:99-114(1991).  
CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT  
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.  
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +  
CC H(+) (out).  
CC -1- SUBCELLULAR LOCATION: Membrane-bound.  
CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X57546; CAA40770.1; -;  
DR PIR: S17994; S17994.  
KM Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.  
FT TRANSMEM POTENTIAL.  
SQ SEQUENCE 48 AA; 5711 MW; 9F5E8A529094F8C3 CRC64;  
  
Query Match 33.3%; Score 22; DB 1; Length 48;  
Best Local Similarity 80.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 KLVPE 5  
Db 3 QLVPE 7  
  
RESULT 30  
ATP8\_SCHPO STANDARD; PRT; 48 AA.  
ID ATP8\_SCHPO  
AC P21536;  
DT 01-MAY-1991 (Rel. 18, Created)

```

DT 15-DEC-1998 (Rel. 37, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE ATP synthase protein 8 (Ec 3.6.3.14) (ATPase subunit 8) (A6L).
GN ATP8.
OS Schizosaccharomyces pombe (Fission yeast).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_taxid:4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:AD7-50;
RA Lang B.F.;
RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) -> ADP + phosphate +
CC H(+) (out).
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X54421; CAA38291.1; -
KM Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT TRANSMEM 8 32 POTENTIAL.
SQ SEQUENCE 48 AA; 5636 MW; 36B3BF480A5FE803 CRC64;

```

Query Match 33.3%; Score 22; DB 1; Length 48;  
Best Local Similarity 80.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 KLVPF 5
   :| | | |
Db 3 QLVPF 7

```

Search completed: July 1, 2003, 17:56:39  
Job time : 14 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2003, 17:54:40 ; Search time 27 Seconds  
(without alignments)  
99.208 Million cell updates/sec

Title: US-09-993-366-1  
Perfect score: 66  
Sequence: 1 KLVFATLHERL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 33835

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database :

SPREMBL\_21:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.podent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	51.5	50	4	Q8TDB0
2	33	50.0	29	16	Q9KCO8
3	32	48.5	39	11	P97632
4	31	47.0	25	2	O06950
5	30	45.5	27	12	O56841
6	29	43.9	31	6	O9XS80
7	28.5	43.2	50	5	O9NM20
8	28	42.4	16	2	O9F9S4
9	28	42.4	17	2	O9ETL3
10	28	42.4	23	2	O9R314
11	28	42.4	36	11	P97644
12	28	42.4	38	4	O9UE53
13	28	42.4	41	6	O9MYR9
14	28	42.4	41	6	O9MZB6
15	28	42.4	41	6	O9MZC4
16	28	42.4	45	10	O49783

17	28	42.4	45	10	O65721	065721 arabidopsis
18	28	42.4	45	13	O9PR01	O9PR01 oncorhynch
19	28	42.4	46	5	O9TXH0	O9TXH0 schistosoma
20	28	42.4	48	11	O9ERC4	O9ERC4 ratius norv
21	28	42.4	49	16	O8U500	O8U500 agrobacteri
22	28	42.4	50	16	O8VK44	O8VK44 mycobacteri
23	27	40.9	27	11	O8VB28	O8VB28 mus musculu
24	27	40.9	28	15	O41600	O41600 human immun
25	27	40.9	36	2	O9K1K7	O9K1K7 arabidopsi
26	27	40.9	37	15	O9RC14	O9RC14 human immun
27	27	40.9	48	2	O54872	O54872 streptococ
28	26	39.4	37	11	O61668	O61668 mus musculu
29	26	39.4	38	5	O9BM37	O9BM37 dugesia tig
30	26	39.4	38	5	O9BM36	O9BM36 dugesia tig
31	26	39.4	39	2	O33073	O33073 mycobacteri
32	26	39.4	47	10	O48633	O48633 hordeum vul
33	26	39.4	50	16	O8ZCH1	O8ZCH1 yersinia pe
34	25	37.9	27	4	O9R5V8	O9R5V8 pseudomonas
35	25	37.9	27	4	O16477	O16477 homo sapien
36	25	37.9	29	8	O9G606	O9G606 gonococphu
37	25	37.9	33	6	O9TS80	O9TS80 sus scrofa
38	25	37.9	36	2	O47528	O47528 escherichia
39	25	37.9	38	5	O9BM35	O9BM35 dugesia tig
40	25	37.9	41	6	O9MYL8	O9MYL8 pan paniscu
41	25	37.9	42	8	O34929	O34929 laeviewceph
42	25	37.9	42	8	O35283	O35283 orocastus p
43	25	37.9	42	8	O36891	O36891 exiltianus e
44	25	37.9	43	4	O9NZ10	O9NZ10 homo sapien
45	25	37.9	48	16	O8X2R9	O8X2R9 escherichia
46	24.5	37.1	25	6	O9N256	O9N256 pan troglod
47	24.5	37.1	25	6	O9N255	O9N255 papio hamad
48	24.5	37.1	25	6	O9N254	O9N254 macaca fasc
49	24.5	37.1	25	6	O9N253	O9N253 macaca mula
50	24	36.4	18	5	O917N8	O917N8 drosophila
51	24	36.4	34	4	O9Y428	O9Y428 homo sapien
52	24	36.4	36	4	O9UMV0	O9UMV0 homo sapien
53	24	36.4	36	11	O92147	O92147 ratius norv
54	24	36.4	37	2	O55239	O55239 synechocyst
55	24	36.4	37	10	O42236	O42236 arabidopsis
56	24	36.4	40	4	O8TE10	O8TE10 homo sapien
57	24	36.4	41	2	O9ZF52	O9ZF52 staphylococ
58	24	36.4	41	4	O9UDG5	O9UDG5 homo sapien
59	24	36.4	41	11	O63974	O63974 mus sp. ais
60	24	36.4	41	16	O8VKR0	O8VKR0 mycobacteri
61	24	36.4	43	10	O9ZUR9	O9ZUR9 arabidopsis
62	24	36.4	43	17	O9HWC8	O9HWC8 halobacteri
63	24	36.4	44	5	O8TF97	O8TF97 aplysia cal
64	24	36.4	44	6	O8SP44	O8SP44 macaca mula
65	24	36.4	44	10	O9LG13	O9LG13 arabidopsis
66	24	36.4	44	16	O8VKJ2	O8VKJ2 mycobacteri
67	24	36.4	46	2	O33589	O33589 staphylococ
68	24	36.4	46	2	O9FI05	O9FI05 staphylococ
69	24	36.4	46	2	O9F623	O9F623 staphylococ
70	24	36.4	46	2	O9L561	O9L561 staphylococ
71	24	36.4	46	2	O53643	O53643 staphylococ
72	24	36.4	47	6	O9TR99	O9TR99 sus scrofa
73	24	36.4	48	6	O9TS86	O9TS86 bos taurus
74	24	36.4	48	9	O9G040	O9G040 bacterioph
75	24	36.4	48	11	O9RIL3	O9RIL3 mus musculu

# ALIGNMENTS

RESULT 1  
ID O8TDB0 PRELIMINARY: PRT: 50 AA.  
AC O8TDB0: 01-JUN-2002 (TREMBUREL. 21, Created)  
DT 01-JUN-2002 (TREMBUREL. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBUREL. 21, Last annotation update)  
DE Apolipoprotein A-1 A175p variant (Fragment).  
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Booth D., Bybee A., Pepsy M., Hawkins P.;
RT "Hereditary amyloidosis.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF485255; AAL92035.1; -.
KW Lipoprotein.
FT NON_TER 1 1
FT NON_TER 50 50
SQ SEQUENCE 50 AA; 5607 MW; A65219AE9E965732 CRC64;

Query Match 51.5%; Score 34; DB 4; Length 50;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LVFATLHERL 13
Db 9 LAPYSDLRRL 20

RESULT 2
O9KC08 PRELIMINARY; PRT; 29 AA.
AC O9KC08;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Hypothetical protein BH1766.
GN BH1766.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE:20512582; PubMed:11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001513; BAB05485.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 29 AA; 3388 MW; EC6A34FC910643CB CRC64;

Query Match 50.0%; Score 33; DB 16; Length 29;
Best Local Similarity 62.5%; Pred. No. 43;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 FATELHER 12
Db 19 FATQLHDK 26

RESULT 3
P97632 PRELIMINARY; PRT; 39 AA.
AC P97632;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Monomeric GTP-binding protein (Fragment).
GN RAB0.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Burcelin R., Rodriguez-Gabin A., Charon M., Almazan G., Larocca J.;
RT "Analysis of the monomeric GTP-binding protein of oligodendrocyte.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U77309; AAB19214.1; -.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4404 MW; 636D405387F7C1C CRC64;

Query Match 48.5%; Score 32; DB 11; Length 39;
Best Local Similarity 53.8%; Pred. No. 90;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13
Db 26 KTVPCGNELHKL 38

RESULT 4
O06950 PRELIMINARY; PRT; 25 AA.
AC O06950;
DT 01-JUL-1997 (TREMblrel. 04, Created)
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE PdxB (Fragment).
GN PDXB.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:97234653; PubMed:9079927;
RA Karlinsey J.E., Pease A.J., Winkler M.E., Bailey J.L., Hughes K.T.;
RT "The flk gene of Salmonella typhimurium couples flagellar P- and L-
RT ring assembly to flagellar morphogenesis.";
RL J. Bacteriol. 179:2389-2400(1997).
DR EMBL; U77618; AAB51500.1; -.
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2909 MW; 8FA652B0F3DD0FEA CRC64;

Query Match 47.0%; Score 31; DB 2; Length 25;
Best Local Similarity 54.5%; Pred. No. 88;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 VVPFATELHERL 13
Db 9 MPYARLFSRL 19

RESULT 5
O56841 PRELIMINARY; PRT; 27 AA.
AC O56841;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Hypothetical 3.1 kDa protein.
OS Vibrio cholerae O139 fsi phage.
OC Viruses.
OX NCBI_TaxID=70203;
RN [1]
RP SEQUENCE FROM N.A.
RA Honma Y.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE:97457190; PubMed:9311128;
RA Bhara M., Shimodori S., Kojima F., Ichinose Y., Hirayama T.,
RA Albert M.J., Supawat K., Honma Y., Iwanaga M., Amako K.;
RT "Characterization of filamentous phage of Vibrio cholerae O139 and
RT O1.";

```

```

RL  FEMS Microbiol. Lett. 154:293-301(1997).
RN  [3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE-98201718; PubMed-9540841;
RT  Honma Y., Ikema M., Toma C., Ebara M., Iwanaga M.;
RT  "Molecular analysis of a filamentous phage (fsl) of Vibrio cholerae
RT  O139."
RL  Blochlm. Biophys. Acta 1362:109-115(1997).
DR  EMBL; D89074; BAA24168.1; -.
KM  Hypothetical protein.
SQ  SEQUENCE 27 AA; 3073 MW; 488864855ADEB05 CRC64;

Query Match
Best Local Similarity 45.5%; Score 30; DB 12; Length 27;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY  3 VPATLHER 12
DB  11 IFFALPSHR 20

RESULT 6
Q9XS80
ID  Q9XS80      PRELIMINARY;      PRT;      31 AA.
AC  Q9XS80;
DT  01-NOV-1999 (TREMBlrel. 12, Created)
DT  01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT  01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE  Androgen receptor (Fragment).
GN  AR.
OS  Sus scrofa (Pig).
OC  Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX  NCBI_TaxId=9823;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-BREED MEISHAN, AND YORKSHIRE;
RA  Selfert J.N., Larsen N.J., Marklund S., Hu Z.L., Rohrer G.A.,
RA  Rothschild M.F.;
RT  "Genetic linkage and Physical Mapping of the Porcine Androgen Receptor
RT  (AR) Gene."
RL  Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF079783; AAD37025.1; -.
KW  Receptor.
FT  NON_TER
FT  NON_TER
SQ  SEQUENCE 31 AA; 3731 MW; 3F7E6C9D56D8FA18 CRC64;

Query Match
Best Local Similarity 43.9%; Score 29; DB 6; Length 31;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY  4 PFATLHE 11
DB  15 PIARELHQ 22

RESULT 7
Q9NM20
ID  Q9NM20      PRELIMINARY;      PRT;      50 AA.
AC  Q9NM20;
DT  01-OCT-2000 (TREMBlrel. 15, Created)
DT  01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT  01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE  Probable glutamate 5-kinase (Fragment).
GN  IM26.272.
OS  Leishmania major.
OC  Eukaryota; Euzoenzoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX  NCBI_TaxId=5664;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-FRIEDLIN;
RA  Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;

```

```

RL  Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AL160493; CAB97890.1; -.
KM  Kinase.
FT  NON_TER
SQ  SEQUENCE 50 AA; 5610 MW; BA8A20834F762926 CRC64;

Query Match
Best Local Similarity 43.2%; Score 28.5; DB 5; Length 50;
Matches 7; Conservative 2; Mismatches 2; Indels 5; Gaps 1;

QY  1 KLVPF-----ATELHE 11
DB  33 KVIPIINENDATALHE 48

RESULT 8
Q9F9S4
ID  Q9F9S4      PRELIMINARY;      PRT;      16 AA.
AC  Q9F9S4;
DT  01-MAR-2001 (TREMBlrel. 16, Created)
DT  01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT  01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE  Methylase HpyI (Fragment).
GN  M.HpyI.
OS  Helicobacter pylori (Campylobacter pylori).
OC  Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC  Helicobacter.
OX  NCBI_TaxId=210;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=FINLAND 9496;
RA  Van Y., Karhukorpi J., Lahti M., Sirvio A., Riiipinen K., Kolho K.-L.,
RA  Lindahl H., Verkasalo M., Fagerholm R., Rautelin H., Karttunen R.;
RT  "Subtypes of the H. pylori Ica gene in Finnish pediatric patients."
RL  Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF176822; AAG13123.1; -.
DR  Interpro: IPR002294; D12N6_mtfase.
DR  Pfam: PF02086; MethyltransferaseD12; 1.
FT  NON_TER
FT  NON_TER
SQ  SEQUENCE 16 AA; 1944 MW; 7306DD11603768A CRC64;

Query Match
Best Local Similarity 42.4%; Score 28; DB 2; Length 16;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY  1 KLVPFATE 8
DB  9 KLIPFIKE 16

RESULT 9
Q9ETL3
ID  Q9ETL3      PRELIMINARY;      PRT;      17 AA.
AC  Q9ETL3;
DT  01-MAR-2001 (TREMBlrel. 16, Created)
DT  01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT  01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE  Methylase HpyI (Fragment).
GN  M.HpyI OR HpyI.
OS  Helicobacter pylori (Campylobacter pylori).
OC  Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC  Helicobacter.
OX  NCBI_TaxId=210;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=FINLAND 18696;
RA  Van Y., Karhukorpi J., Lahti M., Sirvio A., Riiipinen K., Kolho K.-L.,
RA  Lindahl H., Verkasalo M., Fagerholm R., Rautelin H., Karttunen R.;
RT  "Subtypes of the H. pylori Ica gene in Finnish pediatric patients."
RL  Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=7796;

```

RA Yan Y., Karhukorpi J., Lahti M., Sirvio A., Riipinen K., Kolho K.-L.,  
RA Lindahl H., Verrasalo M., Fagerholm R., Rautelin H., Karttunen R.,  
RT "The subtype of Helicobacter pylori Icea gene in the Finnish pediatric  
patients.";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF176823; AAC31325.1; -  
DR EMBL: AF176090; AAC09227.1; -  
DR InterPro: IPR002294; D12N6\_mtfase.  
DR Pfam: PF02086; Methyltransf12; 1.  
FT NON\_TER 17  
SQ SEQUENCE 17 AA: 2058 MW: B2B306D1D1160376 CRC64;

Query Match 42.4%; Score 28; DB 2; Length 17;  
Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KLVPFATE 8  
Db 9 KLIFPKE 16

## RESULT 10

ID 09R314 PRELIMINARY; PRT; 23 AA.  
AC 09R314;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE DNA adenine methylase homolog (Fragment).  
GN M. HpyI.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=219, AND 213;  
RA Raudoniklene A., Berg D.E.;  
RT "Icea2 segment from H. pylori (Alaska strain 219).";  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF008929; AAC64502.1; -  
DR EMBL: AF008928; AAC64500.1; -  
DR InterPro: IPR002294; D12N6\_mtfase.  
DR Pfam: PF02086; Methyltransf12; 1.  
KW Methyltransferase.  
FT NON\_TER 23  
SQ SEQUENCE 23 AA: 2635 MW: 6D6CC63737422B3 CRC64;

Query Match 42.4%; Score 28; DB 2; Length 23;  
Best Local Similarity 62.5%; Pred. No. 3e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KLVPFATE 8  
Db 9 KLIFPKE 16

## RESULT 11

ID P97644 PRELIMINARY; PRT; 36 AA.  
AC P97644;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Zinc finger protein 5 (Fragment).  
GN AGF5.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=96220733; PubMed=8635150;

RA Mesa K., Gebelein B., Cook T., Urrutia R.;  
RT "Identification and characterization of zinc finger encoding genes  
from the tumoral exocrine pancreatic cell line AR42J.";  
RL Cancer Lett. 103:143-149(1996).  
DR EMBL: U78116; AAB36788.1; -  
DR InterPro: IPR000822; ZnF\_C2H2.  
DR Pfam: PF00096; ZF\_C2H2; 1.  
DR SMART: SM00355; ZnF\_C2H2; 1.  
DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 1.  
DR PROSITE: PS0157; ZINC\_FINGER\_C2H2\_2; 1.  
KW DNA-binding; Metal-binding; Zinc-finger.

FT NON\_TER 1  
FT NON\_TER 36  
SQ SEQUENCE 36 AA: 4090 MW: 5F8C8C27B46D7018 CRC64;

Query Match 42.4%; Score 28; DB 11; Length 36;  
Best Local Similarity 45.5%; Pred. No. 4.7e+02;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 LVFPATELHER 12  
Db 10 LVPSLOMHER 20

## RESULT 12

ID 09UE53 PRELIMINARY; PRT; 38 AA.  
AC 09UE53;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Apolipoprotein B-100 gene intron A1 and flanking regions  
DE (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Magender R., Pflitzer R., Stoffel W.;  
RT "Studies on the Organization of the Human Apolipoprotein B-100 gene.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 368:0-0(1987).  
DR EMBL: X04867; CAA28558.1; -  
DR Lipoprotein.  
KW Lipoprotein.  
FT NON\_TER 1  
FT NON\_TER 38  
SQ SEQUENCE 38 AA: 4634 MW: BBSDE6D401437EC7 CRC64;

Query Match 42.4%; Score 28; DB 4; Length 38;  
Best Local Similarity 58.3%; Pred. No. 5e+02;  
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 KLVPFATELHER 12  
Db 24 KLTFKTELRRVR 35

## RESULT 13

ID 09MYL9 PRELIMINARY; PRT; 41 AA.  
AC 09MYL9;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE Apolipoprotein B (Fragment).  
GN APOB.  
OS Pan troglodytes schweinfurthii.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.  
OX NCBI\_TaxID=37010;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=17, 18, AND 12;

```

RX MEDLINE=20086521; PubMed=10618587;
RA Delnard A.S., Kidd K.;
RT "Identifying conservation units within captive chimpanzee
  populations."
RL Am. J. Phys. Anthropol. 111:25-44(2000).
DR EMBL: AF245203; AAF75256.1; -.
DR EMBL: AF245189; AAF75242.1; -.
DR EMBL: AF245198; AAF75251.1; -.
KW Lipoprotein.
FT NON_TER 1 1
FT SEQUENCE 41 AA; 4943 MW; D9EE25FCDA6935E3 CRC64;
SQ
Query Match 42.4%; Score 28; DB 6; Length 41;
Best Local Similarity 58.3%; Pred. No. 5.4e+02;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVFATELHER 12
   ||| |||
Db 19 KLIFFTELRLVR 30

RESULT 14
Q9MZB6 PRELIMINARY; PRT; 41 AA.
ID Q9MZB6; Q9MZB7; Q9MZB8; Q9MZB9; Q9MZC0; Q9MZC1; Q9MZC2; Q9MZC3;
AC Q9MZC5; Q9MZC6; Q9MZC7; Q9MZC8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Apolipoprotein B (Fragment).
GN APOB.
OS Pan troglodytes troglodytes.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OC NCBI_TaxID=37011;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16, 15, 14, 13, 11, 10, 9, 8, 6, 5, 4, AND 3;
RA MEDLINE=20086521; PubMed=10618587;
RT Delnard A.S., Kidd K.;
RT "Identifying conservation units within captive chimpanzee
  populations."
RL Am. J. Phys. Anthropol. 111:25-44(2000).
DR EMBL: AF245202; AAF75255.1; -.
DR EMBL: AF245201; AAF75254.1; -.
DR EMBL: AF245200; AAF75253.1; -.
DR EMBL: AF245199; AAF75252.1; -.
DR EMBL: AF245197; AAF75250.1; -.
DR EMBL: AF245196; AAF75249.1; -.
DR EMBL: AF245195; AAF75248.1; -.
DR EMBL: AF245194; AAF75247.1; -.
DR EMBL: AF245192; AAF75245.1; -.
DR EMBL: AF245191; AAF75244.1; -.
DR EMBL: AF245190; AAF75243.1; -.
DR EMBL: AF245188; AAF75241.1; -.
KW Lipoprotein.
FT NON_TER 1 1
FT SEQUENCE 41 AA; 4943 MW; D9EE25FCDA6935E3 CRC64;
SQ
Query Match 42.4%; Score 28; DB 6; Length 41;
Best Local Similarity 58.3%; Pred. No. 5.4e+02;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVFATELHER 12
   ||| |||
Db 19 KLIFFTELRLVR 30

RESULT 15
Q9MZC4 PRELIMINARY; PRT; 41 AA.
ID Q9MZC4

```

```

AC Q9MZC4; Q9MZC9; Q9MZD0;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Apolipoprotein B (Fragment).
GN APOB.
OS Pan troglodytes troglodytes.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OC NCBI_TaxID=37012;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7, 2, AND 1;
RA MEDLINE=20086521; PubMed=10618587;
RT Delnard A.S., Kidd K.;
RT "Identifying conservation units within captive chimpanzee
  populations."
RL Am. J. Phys. Anthropol. 111:25-44(2000).
DR EMBL: AF245193; AAF75246.1; -.
DR EMBL: AF245187; AAF75240.1; -.
DR EMBL: AF245186; AAF75239.1; -.
KW Lipoprotein.
FT NON_TER 1 1
FT SEQUENCE 41 AA; 4943 MW; D9EE25FCDA6935E3 CRC64;
SQ
Query Match 42.4%; Score 28; DB 6; Length 41;
Best Local Similarity 58.3%; Pred. No. 5.4e+02;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVFATELHER 12
   ||| |||
Db 19 KLIFFTELRLVR 30

RESULT 16
Q49783 PRELIMINARY; PRT; 45 AA.
ID Q49783
AC Q49783;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE R2R3-MYB transcription factor (Fragment).
GN ATMYS52.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Romero I., Fuentes A., Benito M.J., Malpica J., Leyva A., Paz-Ares J.;
RT "One hundred R2R3-MYB genes in the genome of Arabidopsis thaliana."
RT Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL: Z95775; CAB09207.1; -.
DR HSSP: Q03237; IASU.
DR InterPro: IPR001005; MYB_DNA_binding.
DR Pfam: PF00249; myb_DNA_binding; 1.
DR SMART: SM00395; SANT; 1.
DR PROSITE: PS50090; MYB_3; 1.
KW DNA-binding; Nuclear protein.
FT NON_TER 1 1
FT SEQUENCE 45 AA; 5562 MW; 087262E9A860855F CRC64;
SQ
Query Match 42.4%; Score 28; DB 10; Length 45;
Best Local Similarity 60.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 PFATELHERL 13
   ||| |||

```

Db 15 PFTEEEERL 24

## RESULT 17

ID 065721 PRELIMINARY: PRT: 45 AA.

AC 065721.

DT 01-AUG-1998 (TrEMBLrel. 07, Created)

DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE AtMYB54 protein (Fragment).

GN AtMYB54.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

DR HSSP: P06876; 1IDY.

DR InterPro: IPR001005; MYB\_DNA\_binding.

DR Pfam: PF00249; myb\_DNA-binding; 1.

DR SMART: SM00395; SANT; 1.

DR PROSITE: PS00090; MYB\_3; 1.

KW DNA-binding; Nuclear protein.

FT NON\_TER 1

SQ SEQUENCE 45 AA: 5565 MW: 10D16E5FB680855F CRC64;

Query Match 42.4%; Score 28; DB 10; Length 45;

Best Local Similarity 60.0%; Pred. No. 5.9e+02;

Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 PFATELHERL 13

Db 15 PFTEEEERL 24

## RESULT 18

ID 09PT01 PRELIMINARY: PRT: 45 AA.

AC 09PT01.

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Apolipoprotein E (Fragment).

GN APOE.

OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

OC Proteoanchopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

OC NCBI\_TaxID=8022;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=RED BLOOD CELLS.

RX MEDLINE=20098531; PubMed=10632725;

RA Durilat M., Andre M., Babin P.J.;

RT "Conserved protein motifs and structural organization of a fish gene homologous to mammalian apolipoprotein E.",

RL Eur. J. Biochem. 267:549-559(2000).

DR EMBL: AJ132621; CAB65321.1; -.

DR InterPro: IPR000074; Apolipoprotein.

DR Pfam: PF01442; Apolipoprotein; 1.

KW Lipoprotein.

FT NON\_TER 1

SQ SEQUENCE 45 AA: 5082 MW: 9C5E4B0A5D8E522 CRC64;

Query Match 42.4%; Score 28; DB 13; Length 45;

Best Local Similarity 62.5%; Pred. No. 5.9e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVPPATE 8

Db 24 KLVPPAAD 31

## RESULT 19

ID 09TXH0 PRELIMINARY: PRT: 46 AA.

AC 09TXH0.

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Triose-phosphate ISOMERASE-28-kDa AG (Fragments).

OS Schistosoma mansoni (Blood fluke).

OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;

CC Schistosomatidae; Schistosomatidae; Schistosoma.

NCBI\_TaxID=6183;

RN [1]

RP SEQUENCE.

RX MEDLINE=92105755; PubMed=1729373;

RA Harn D.A., Gu W., Oligino L.D., Mitsuyama M., Gebremichael A.,

RA Richter D.;

RT "A protective monoclonal antibody specifically recognizes and alters the catalytic activity of schistosoma triose phosphate isomerase.",

RL J. Immunol. 148:562-567(1992).

DR HSSP: P00938; 1HTI.

DR InterPro: IPR000652; Triophos\_ismrse.

DR PRODOM: PD001005; Triophos\_ismrse; 1.

FT NON\_TER 1

FT NON\_CONS 15

FT NON\_CONS 27

FT NON\_TER 46

SQ SEQUENCE 46 AA: 5034 MW: 3BDE4394B4B703DB CRC64;

Query Match 42.4%; Score 28; DB 5; Length 46;

Best Local Similarity 62.5%; Pred. No. 6.1e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 ATELHERL 13

Db 19 AQEVHEKL 26

## RESULT 20

ID 09ERC4 PRELIMINARY: PRT: 48 AA.

AC 09ERC4.

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE GTP cyclohydrolase I (Fragment).

OC Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY;

RA Shimaji M., Hirayama K., Kapatos G.;

RT "Rat GTP cyclohydrolase I intron 2.",

RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF317087; AAC30952.1; -.

DR HSSP: P27511; 1GTP.

DR InterPro: IPR001474; GTP\_cyclohydrol.

DR Pfam: PF01227; GTP\_cyclohydrol; 1.

DR PRODOM: PD003330; GTP\_cyclohydrol; 1.

DR PROSITE: PS00859; GTP\_CYCLOHYDROL\_1; 1.

KW Hydrolyase.

FT NON\_TER 1 1  
RT NON\_TER 48 48  
SQ SEQUENCE 48 AA; 5580 MW; F6EA44D5152CD5FE CRC64;

Query Match 42.4%; Score 28; DB 11; Length 48;  
Best Local Similarity 55.6%; Pred. No. 6.3e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LVPEATELH 10  
DB 30 LVPEVGRVH 38

## RESULT 21

Q805U0 PRELIMINARY; PRT; 49 AA.

AC Q805U0; 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE AGR\_PAT\_456.  
GN AGR\_PAT\_456.  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
NCBI\_TaxID=176299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21608551; PubMed=11743194;  
RA Goodner B., Hinkle G., Gatlung S., Miller N., Blanchard M.,  
RA Goutello B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
RA Houniel K., Gordon J., Vaudin M., Tarchouk O., Epp A., Liu F.,  
RA Wolanin C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,  
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,  
RA Cielo C., Slater S.;  
RT "Genome sequence of the plant pathogen and biotechnology agent  
RT Agrobacterium tumefaciens C58.";  
RL Science 294:2323-2328(2001).  
DR EMBL; AE007901; AAK90693.1; -.  
KW Plasmid.  
SQ SEQUENCE 49 AA; 5699 MW; 6F5E194169425EAC CRC64;

Query Match 42.4%; Score 28; DB 16; Length 49;  
Best Local Similarity 55.6%; Pred. No. 6.5e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 FATELHERL 13  
DB 34 FAEEYHDL 42

## RESULT 22

Q8VK44 PRELIMINARY; PRT; 50 AA.

AC Q8VK44; 01-MAR-2002 (TReMBLrel. 20, Created)  
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE Hypothetical protein MT1288.  
GN MT1288.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=CDC 1551 / OSHKOSH;  
RA Flaischmann R.D., Allard D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kojanov J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA DeCher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RT Laboratory strains.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE007004; AAK45546.1; -.  
DR TIGR; MT1288; -.  
KW Hypothetical protein.

Query Match 42.4%; Score 28; DB 16; Length 50;  
Best Local Similarity 31.6%; Pred. No. 6.6e+02;  
Matches 6; Conservative 4; Mismatches 3; Indels 6; Gaps 1;

QY 1 KLVEF-----ATELHERL 13  
DB 26 RIPLRAVNRATETTHARI 44

## RESULT 23

Q8VBZ8 PRELIMINARY; PRT; 27 AA.

AC Q8VBZ8; 01-MAR-2002 (TReMBLrel. 20, Created)  
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE Hypothetical 3.0 kDa protein (Fragment).  
OS Mus musculus (mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=LIVER.  
RA Strausberg R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022155; AAH22155.1; -.  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 27 AA; 2974 MW; 56BB3923CE0C7417 CRC64;

Query Match 40.9%; Score 27; DB 11; Length 27;  
Best Local Similarity 55.6%; Pred. No. 5.4e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LVPEATELH 10  
DB 8 LWPFSVSLH 16

## RESULT 24

Q41600 PRELIMINARY; PRT; 28 AA.

AC Q41600; 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=C20;  
MEDLINE=98105804; PubMed=9445059;  
RA Connor R.I., Korber B.T.M., Graham B.S., Hahn B.H., Ho D.D.,  
RA Walker B.D., Neumann A.U., Vermond S.H., Mestecky J., Jackson S.,  
RA Fenimore E., Cao Y., Gao F., Kalam S., Kunstman K.J., McDonald D.,  
RA McWilliams N., Trkola A., Moore J.P., Wolinsky S.M.,  
RT "Immunological and virological analyses of persons infected by human  
RT immunodeficiency virus type 1 while participating in trials of  
RT recombinant gp120 subunit vaccines.";  
RL J. Virol. 72:1552-1576(1998).  
EMBL; U84846; AAC58888.1; -.

KW AIDS: Coat protein; Glycoprotein.

FT NON\_TER 1  
SQ SEQUENCE 28 AA: 3163 MW: 5AE25FD225B16B2 CRC64;

Query Match  
Best Local Similarity 40.9%; Score 27; DB 15; Length 28;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVPFATELH 10  
Db 5 KLTPLCVTLH 14

RESULT 25

Q9KJK7 PRELIMINARY; PRT; 36 AA.

AC Q9KJK7;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE N1FE (Fragment).  
GN N1FE.

OS Anabaena azollae.  
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.  
OX NCBI\_TaxID=1164;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-1A;

RA Dominic B., Zani S., Chen Y.-B., Mellon M.T., Zehr J.P.;

RT "Organization of the nif genes of the nonheterocystous cyanobacterium

RT Trichodesmium sp. IMS101."

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF167540; AAF82650.1;

DR InterPro; IPR000510; Oxred\_nitrognasel.

DR Pfam; PF00148; oxidored\_nitro; 1.

FT NON\_TER 1  
SQ SEQUENCE 36 AA: 4351 MW: 83FC1A3702316442 CRC64;

Query Match  
Best Local Similarity 40.9%; Score 27; DB 2; Length 36;  
Matches 6; Conservative 50.0%; Pred. No. 7.3e+02;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 LVPFATELHERL 13  
Db 10 MVEARELYEAL 21

RESULT 26

Q99C14 PRELIMINARY; PRT; 37 AA.

AC Q99C14;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE Truncated envelope glycoprotein (Fragment).  
GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-C-GC2-1813;

RA MEDLINE=20570520; Pubmed=11121058;

RA Gratton S., Cheynier R., Dumaurier M.J., Oksenhendler E.,

RA Main-Hobson S.;

RT "Highly restricted spread of HIV-1 and multiply infected cells within

RT splenic germinal centers."

RL Proc. Natl. Acad. Sci. U.S.A. 97:14566-14571(2000).

DR EMBL; AF319933; AAG52872.1;

KW Envelope protein.  
FT NON\_TER 1  
SQ SEQUENCE 37 AA: 3937 MW: E41F9C99E878C7E CRC64;

Query Match  
Best Local Similarity 40.9%; Score 27; DB 15; Length 37;  
Matches 5; Conservative 50.0%; Pred. No. 7.5e+02;  
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLVPFATELH 10  
Db 1 KLTPLCVTLH 10

RESULT 27

Q54872 PRELIMINARY; PRT; 48 AA.

AC Q54872;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Hypothetical 5.6 kDa protein (Fragment).  
OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;

OC Streptococcaceae; Streptococcus.

OX NCBI\_TaxID=1313;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=9008767; Pubmed=2676973;

RA Prudhomme M., Martin B., Mejean V., Claverys J.P.;

RT "Nucleotide sequence of the Streptococcus pneumoniae hexB mismatch

RT repair gene: homology of hexB to MuL of Salmonella typhimurium and to

RT PMS1 of Saccharomyces cerevisiae."

RL J. Bacteriol. 171:5332-5338(1989).

DR EMBL; M29686; AAA88599.1;

FT NON\_TER 48  
SQ SEQUENCE 48 AA: 5630 MW: 01299BDA3022B6CD CRC64;

Query Match  
Best Local Similarity 40.9%; Score 27; DB 2; Length 48;  
Matches 5; Conservative 41.7%; Pred. No. 9.8e+02;  
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KLVPFATELHER 12  
Db 29 KVIFAEINDOR 40

RESULT 28

Q61668 PRELIMINARY; PRT; 37 AA.

AC Q61668;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Gene fragment for histone H2a (Fragment).  
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=84041477; Pubmed=6314253;

RA Sittman D.B., Graves R.A., Marzluff W.F.;

RT "Structure of a cluster of mouse histone genes."

RL Nucleic Acids Res. 11:6679-6697(1983).

DR EMBL; X02622; CAA26476.1;

DR InterPro; IPR004822; Histone\_core.

DR Pfam; PF00125; histone; 1.

FT NON\_TER 1  
SQ SEQUENCE 37 AA: 3985 MW: 8BE9A5424140D358 CRC64;

Query Match  
Best Local Similarity 39.4%; Score 26; DB 11; Length 37;  
Matches 5; Conservative 45.5%; Pred. No. 1.2e+03;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 LVPFATELHER 12  
Db 1 LVPFATELHER 12



Db 23 LLPKTESHHK 33

## RESULT 29

Q9BM37 ID Q9BM37 PRELIMINARY; PRT; 38 AA.  
 AC Q9BM37;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE LINE-like reverse transcriptase (Fragment).  
 OS Dugesia tigrida (Planarian).  
 OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;  
 OC Paludicola; Dugesidae; Girardia.  
 OX NCBI\_TaxID=6162;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TRANSPOSON-LRT-L1 RETROTRANSPOSON;  
 RX MEDLINE-20570504; PubMed-11121049;  
 RA Arkhipova I., Meselson M.;  
 RT "Transposable elements in sexual and asexual taxa.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).  
 DR EMBL; AY013959; AAC59941.1; -;  
 KW RNA-directed DNA polymerase.  
 FT NON\_TER 1  
 FT NON\_TER 38  
 SQ SEQUENCE 38 AA; 4141 MW; 838096460C4D719A CRC64;

## Query Match

Best Local Similarity 39.4%; Score 26; DB 5; Length 38;  
 Pred. No. 1.2e+03;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 PFATELH 10  
 | ||:|  
 Db 14 PILTEIH 20

## RESULT 30

Q9BM36 ID Q9BM36 PRELIMINARY; PRT; 38 AA.  
 AC Q9BM36;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE LINE-like reverse transcriptase (Fragment).  
 OS Dugesia tigrida (Planarian).  
 OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;  
 OC Paludicola; Dugesidae; Girardia.  
 OX NCBI\_TaxID=6162;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TRANSPOSON-LRT-L2 RETROTRANSPOSON;  
 RX MEDLINE-20570504; PubMed-11121049;  
 RA Arkhipova I., Meselson M.;  
 RT "Transposable elements in sexual and asexual taxa.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).  
 DR EMBL; AY013960; AAC59942.1; -;  
 KW RNA-directed DNA polymerase.  
 FT NON\_TER 1  
 FT NON\_TER 38  
 SQ SEQUENCE 38 AA; 4115 MW; 838096549C4D719A CRC64;

## Query Match

Best Local Similarity 39.4%; Score 26; DB 5; Length 38;  
 Pred. No. 1.2e+03;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 PFATELH 10  
 | ||:|  
 Db 14 PILTEIH 20

Search completed: July 1, 2003, 17:57:14  
 Job time : 30 secs

**This Page Blank (uspto)**